PL-TR-95-2036

IMPROVED ESTIMATION OF THE EARTH'S GRAVITY FIELD

Richard H. Rapp
Christopher Jekeli

The Ohio State University
Department of Geodetic Science and Surveying
Columbus, OH 43210

April 1995

Final Report
22 October 1993 - 22 April 1995

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

PHILLIPS LABORATORY
Directorate of Geophysics
AIR FORCE MATERIEL COMMAND
HANSCOM AIR FORCE BASE, MA 01731-3010

"This technical report has been reviewed and is approved for publication"

(Signature)
Anestis J. Romaides
Contract Manager

(Signature)
David N. Anderson
Chief, Ionospheric Modelling &
Remote Sensing Branch

(Signature)
William K. Vickery
Director, Ionospheric Effects Division

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
April 1995

3. REPORT TYPE AND DATES COVERED
Final Report 10/22/93 - 4/22/95

4. TITLE AND SUBTITLE
Improved Estimation of the Earth's Gravity Field

5. FUNDING NUMBERS
F19628-93-K-0033
PE 663200
PR 3204 TA DM
WU AP

6. AUTHOR(S)
Richard H. Rapp
Christopher Jekeli

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Department of Geodetic Science and Surveying
The Ohio State University
Columbus, OH 43210

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING, MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Phillips Laboratory
29 Randolph Rd
Hanscom Air Force Base, MA 01731-3010
Contract Manager: Anestis Romaides / GPIM

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
PL-TR-95-2036

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
This final report describes research activities carried out by three investigators. The studies of N. Balsubramania related to vertical datums and how one might use geodetic techniques to connect the many vertical datums of the world. In addition he implemented a data analysis using information at 17 space geodetic stations in six regional vertical datums. This data, with gravity field information, was used to define a global vertical datum and relate the regional datums to the ideal datum.
Rapp prepared several papers related to the vertical datum question. The papers emphasized the importance of the geoid undulation, or height anomaly, in connecting vertical datums. Aspects of this research led to a proposal for a World Vertical Datum. In order to determine an improved gravity field for vertical datum purpose, a joint DMA/GSFC project was established. This report describes the role that Rapp and Jekeli have played in the project.
Jekeli pursued studies related to improved ways to carry out a harmonic analysis of discrete and average values. These studies developed equations to describe the aliasing error of the harmonic coefficient estimate. Equations are described that, when used to calculate average values, reduce the aliasing effects. Two papers and two presentations reflecting research carried out under this contract are included as an appendix.

14. SUBJECT TERMS
Gravity, Geoid, Height Ststensm Vertical Datums

15. NUMBER OF PAGES
56

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	SAR

TABLE OF CONTENTS

1. Introduction

This report is the final report of the Phillips Laboratory contract F19628-93-K-0033. This contract was funded in response to a proposal that was submitted to Phillips Laboratory in July 1993 with the title Improved Estimation of the Earth's Gravity Field. The starting date of the contract was 22 October 1993 with an ending date of 21 January 95. On May 31, 1994 a request was made to Anestis Romaides to expand the scope of the original proposal with additional funding requested. In July 1994 the modification of the contract was approved and the contract ending data was extended to 22 April 1995.

2. Research Activities

The proposal submitted in July 1993 described the importance of improving our knowledge of the Earth's gravity field and its use in the ultimate development of a world height system. In the context of the original proposal the Earth's gravity field, or more properly the Earth's gravitational potential was to be represented by a set of potential coefficients and techniques for the improved estimation of these coefficients was part of this project's area of interest. In the May 31, 1994 letter to Mr. Romaides the work of the project was to be expanded so that project personnel could participate in the research and meetings related to the joint gravity and geoid improvement project of the Defense Mapping Agency and the NASA, Goddard Space Flight Center. This report has been written to describe the activities, undertaken by project personnel, consistent with the goals defined by the proposals. The persons involved with research supported by this project are Richard H. Rapp, Professor Emeritus, Christopher Jekeli, Associate Professor, Nagarajan Balasubramania, Dru Smith Graduate Research Associates. For convenience, this final report has been divided into sections according to the author.

2.1 Research Activities of N. Balasubramania

Mr. Balasubramania continued his activities related to the development of a world height system. His research led to the publication of Scientific Report No. 1, in April 1994, "Definition and Realization of a Global Vertical Datum," PL-TR-94-2144,ADA283303, the abstract of this report is as follows:

Definition and Realization of A Global Vertical Datum

Nagarajan Balasubramania

Abstract

The requirements for defining and realizing a global vertical datum (GVD) with a precision on the dm-or even cm-level has become essential now, due to the improved measurement accuracy provided by the space geodetic techniques. This study identifies the different approaches in which such a global vertical datum could be evolved.

The first approach, which can be considered as an ideal approach, uses the best available geocentric station positions from present and future space geodetic networks, highly accurate geopotential model and surface gravity data around the space stations for defining and realizing a global vertical datum. The other simplified approach, which can be considered more of an operational development with a short time framework, is mainly dependent on the widely available GPS and DORIS tracking networks and accurate geoid height models for the GVD development.

After a review of the various heights and height systems that are in use today, detailed mathematical procedures for setting up the observation equations to define a global vertical datum in a least-squares sense, under different data scenarios, are discussed. A test computation to realize the first iteration global vertical datum with available data at 17 space geodetic stations in six regional vertical datums is also included. The gravimetric height anomaly/undulation computations at the space geodetic stations required in the test computation have been performed using both Modified Stokes' technique and least squares collocation technique combining surface gravity data in a small cap around the stations with potential coefficients from OSU91A model. The test results show that in an idealistic approach the global vertical datum can be realized to an accuracy of ±5cm an its connection to the regional vertical datums to an accuracy of ±5cm to ±23cm. The estimated values of separation between different regional vertical datums agree mostly well with the results reported by various geodesists and oceanographers based on their regional studies.

The Balasubramania report also appeared as Report 427 of the Department of Geodetic Science and Surveying, April 1994. The report was also submitted to the Graduate School of The Ohio State University in partial fulfillment of the requirements for the Degree Doctor of Philosophy. Basasubramania received his PhD from Ohio State in March 1994.

2.2 Research Activities of Richard H. Rapp

During this time period Professor Rapp served as the adviser to Balasubramania and carried out activities related to geoid and gravity model improvement.

The interest in vertical datums was continued in this project. In December 1993 made two presentations in Wormley, England, related to project interests. The first presentation was: A Global Vertical Reference Frame. This was at the special meeting of the group involved with the Geodetic Fixing of Tide Gauge Bench Marks." The meeting of this group was followed by the meeting of the International Association of Geodesy Special Study Group 5.149, Vertical Datum Investigation. At this meeting Rapp made the presentation: A World Vertical Datum Proposal. Copies of the viewgraphs used in this talk are given in the Appendix of this report.

Upon invitation, Rapp prepared the following paper for presentation at the meeting for the International Symposium on Marine Positioning, Hannover, Germany: Vertical Datums in Land and Marine Positioning. This symposium was held in September 1994. The complete text of this paper is contained in the Appendix of this report. The paper was also published in the Proceedings of the Symposium published by the Marine Technology Society, Washington, D.C.

Rapp has also been involved in the joint DMA/GSFC project for gravity field and geoid improvement. He is the chair of the steering committee and participates in the activities of the working groups charged with the implementation of the procedures that will lead to a significantly improved degree 360 model. During this contract period, Rapp chaired two meetings (April 5 and November 1, 1994) of the steering group. He prepared minutes of these meetings which were distributed to all parties of the joint effort. In addition to the steering group meetings Rapp attended working group meetings held 8-9 June 94 (NASA/GSFC) and 1-2 August 94 (DMAAC). At the July meeting Rapp made a presentation on: Terrain Base (NGDC) and Ice Thickness Information. This informal presentation described comparisons made with several different elevation files. Specific comparisons were carried out by Dru Smith, Graduate Research Associate, in a 1°x1° cell (NW corner, $\phi=20°$, $\lambda=204°$), in the Hawaiian area. 5'x5' elevations were computed with a 3"x3" DTED file provided to Ohio several years ago. These mean elevations were then compared to the elevations in three data bases: ET0P05U, T Base (α version) and T Base (β version). The Terrain

Base data sets were provided to us by NGDC for evaluation. Statistics, as complied by Smith, are given in Table 1

Table 1
Comparison of 5'x5' Mean Elevations From Three Data Sets to Ground Truth Values Estimated from 3"x3" Values

	Model		
	ET0P05U	T Base (α)	T Base (β)
No. of Comp	89	87	87
Mean Diff.	530m	475m	155m
Std. Dev.	±916m	±718m	±434m
RMS Diff	1058m	861m	461m
Min Diff	-1217m	-730m	-730m
Max Diff	2548m	2936m	1499m
No. of Diff where $\|D\|>100m$	37	18	5

The conclusion from this test was that Terrain Base (β) was, in this area, a significant improvement over the older ET0P05U and the initial version of Terrain Base.

Rapp and R.S. Nerem of NASA/GSFC, prepared a paper for the Joint Symposium of the International Gravity Commission and the International Geoid Commission that was held in Graz, Austria, September 11-17, 1994. The title and abstract of this paper are as follows:

A JOINT GSFC/DMA PROJECT FOR IMPROVING THE MODEL OF THE EARTH'S GRAVITATIONAL FIELD

Richard H. Rapp
Department of Geodetic Science and Surveying
The Ohio State University
Columbus, Ohio 43210

R. Steven Nerem
NASA Goddard Space Flight Center
Space Geodesy Branch, Code 926
Greenbelt, MD 20771

ABSTRACT

The U.S. Defense Mapping Agency and the NASA Goddard Space Flight Center with the aid of other organizations such as The Ohio State University are cooperating in a joint effort to determine a significantly improved degree 360

spherical harmonic model representing the Earth's gravitational potential. This new model will be of immediate use in defining a geoid undulation model that will be the basis for an enhanced WGS84 geoid.

The development of the new model is driven, in part, by the need to determine an accurate geoid undulation model that will be the reference surface for a World Height System to be implemented in the 1996 time period. In addition, the new geoid estimation will help satisfy increasingly important studies in ocean circulation (sea surface topography) and geodetic positioning through GPS.

The new model estimation will incorporate existing and new satellite data. New data will include GPS tracking of Topex/Poseidon, Doris tracking of several satellites, altimeter data from Topex/Poseidon, and ERS-1 and Doppler data from satellites at inclinations not covered or weakly represented in previous solutions.

The surface gravity data to be used in the new solution will be based on an updated 30' mean anomaly data base developed at the DMA Aerospace Center. This new data set will incorporate a substantial amount of new data that has, and will, become available in Europe, the FSU, South America, Greenland, Africa, Asia and Antarctica. Anomaly values in areas such as Canada, the United States and Australia will be based on updated data files. This data will be used, after suitable corrections, to form normal equations that can be used with the satellite derived normal equations.

In addition, 30'x30' mean anomalies derived from the Geosat Geodetic Mission satellite altimeter data will be used in the project. The file will be merged with the files based on the surface terrestrial data. In areas where no data exists, anomaly estimates will be made from new elevation data through topographic isostatic models to ultimately yield a global 30' anomaly file. In addition, an updated 1°x1° anomaly file based on terrestrial data (both land and ocean) will be determined.

The final stage of the data processing will be the development of several degree 360 models using different data sets and weighting procedures. The current plan is to use existing software, for the combination solution, with minimum modifications to assure a timely effort. Several preliminary models will be made available to the international community for evaluation. A final model will be selected based on extensive tests of the preliminary models. The final model and accuracy estimates will be released in mid 1996. The model will be used to determine accurate geoid undulations that will be available in gridded form.

This paper will be published in the proceedings of the meeting by Springer Verlag in the IAG series. Copies of the complete paper were provided to Phillips Laboratory and DMA. The paper was to have been presented by Rapp but due to unforeseen circumstances he was not able to attend the meeting. The paper was delivered by Dr. M. Kumar of DMA.

Additional studies were carried out to consider precise procedures for the determination of geoid undulations from the potential coefficient representation of the Earth's gravitational potential. Previous discussion on this evolved knowing the difficulty in evaluating the disturbing potential at a surface (the geoid) interior to the Earth. An alternative suggestion has been made that the disturbing potential be evaluated at the external surface of the Earth and then downward continued to the geoid using additional information. The downward continuation can be done using equations developed many years ago by Moritz who showed the relationship between the geoid undulation and height anomaly. This relationship depends on Bouguer gravity anomalies that can be computed from a potential coefficient model if topographic elevations are known. Postulating that a spherical harmonic expansion of terrain elevations can be made procedures can be developed to represent the correction term to convert height anomalies to geoid undulations. In a simplified case correction terms for the potential coefficient can be computed. This could lead to the determination of a modified set of potential coefficients which could be used for geoid undulation computation using traditional software. The problem is complicated by the fact that evaluation of the disturbing

potential at altitude requires knowledge of the terrain elevation. This significantly reduces the efficiency of computer software designed for the fast calculation of geoid undulations, from potential coefficient models, on a grid. The best way to carry out these computations has not yet been determined.

A review of methods for the calculation of geoid undulations from potential coefficient models was prepared by Rapp for this contract and for presentation at the International School for the Determination and Use of the Geoid. This school was held in Milan, Italy, October 1994, under the auspices of Professor Fernandó Sansó, Director. The title of this paper is; The Use of Potential Coefficient Models in Computing Geoid Undulations. Copies of the paper have been provided to Phillips Laboratory and the Defense Mapping Agency. The paper will be published in the bulletin series of the International Geoid Service in 1995.

2.3 Research Activities of Christopher Jekeli

The traditional method of spherical harmonic analysis of discrete, global gravity anomaly data was popularized by Colombo who reformulated the technique for FFT utiliziation, thus offering the ability to perform very high degree expansions with minimal computational effort. However, the method is still based on a discretization of the integral (a quadrature formula) that defines the spherical harmonic coefficient in terms of the anomaly data. Other methods based on least-squares collocation and least-squares adjustment (also developed by Colombo) similarly discretized the integral (albeit with different weights), a necessity because the data are discrete.

Each of these methods, then, is subject to an error, known in the parlance of spectral analysis, as an aliasing error. It is an error independent of the noise in the data, but does depend on the amount of information contained in the data at frequencies generally higher than the frequency defined by the sampling interval (the Nyquist frequency). In theory, and if certain assumptions of the geo-statistical nature of the gravity field are fulfilled, the least-squares collocation technique should yield the best (minimum aliasing error) estimate of the harmonic coefficients. The aliasing error can be estimated empirically using simulation studies; and some simple models were proposed for the OSU89 and previous models. This, in turn, led to simplified weighting schemes for potentially better estimation of the harmonic coefficients.

This rather ad hoc approach to harmonic analysis is unsatisfactory and calls for a more systematic characterization of the aliasing error. This was accomplished by Jekeli who derived explicit formulas for the aliasing error for the traditional quadratures formula, and by analogy for the least-squares adjustment formula. Also, a modified estimation was proposed that reduces the aliasing error by determining twice the number of harmonic coefficients (thus, the same number as number of observations) at only slightly greater computational cost.

In addition, by clearly understanding the origin of the aliasing error in the analysis of data on a sphere, it became apparent that certain types of data smoothing are better than others in reducing the aliasing error. Typically, data smoothing is the result of uniformly weighted averaging of gravity anomalies over blocks defined by latitude and longitude lines. The aliasing error formulas show how this procedure fails to attenuate a part of the high-degree spectrum that then aliases the estimated spectrum. Averages over constant spherical caps, though spoiled at higher latitudes by the increasing correlation between neighbors on the latitude/longitude grid, are more definitive in filtering out the high-degree components. Special weighting functions (such as the Gaussian function) can further reduce the aliasing that remains due to the ringing of the uniformly weighted average.

These results were presented (Jekeli, 1994) and accepted for publication (Jekeli, 1995) with the following summary. Presented material and a preprint of the submitted paper are included in the Appendix to this report.

Summary
The methods of harmonic analysis on the sphere were investigated from the point of view of aliasing. The aliasing error was formulated in terms of spherical harmonic coefficients in the case when the function is sampled on a regular grid of latitudes and longitudes. This yielded several results, some of which are known, but perhaps more analytically illustrated, here: 1) The simple quadratures method and related methods are biased even with band-limited functions. 2) A new method that eliminates this bias is also superior to Colombo's method of least squares in terms of reducing aliasing, because it estimates more coefficients at little added computational cost. 3) But, a simple modification of the model makes the least squares method identical to the new method as one is the dual of the other. The new method solves for as many coefficients as data and, in the presence of data noise, it need not use fabricated noise covariances to make the solution numerically tractable. 4) The essential elimination of aliasing can only be effected with spherical cap averages, not with the often used constant angular block averages.

3. Personnel

Persons who carried out activities under this contract were: Richard H. Rapp, Christopher Jekeli, Nagarajan Balasubramania, and Dru Smith. Professor Rapp and Jekeli are professors in the Department of Geodetic Science and Surveying while Dr. Balasubramania and Mr. Smith were Graduate Research Associates.

4. Travel

During this contract Professor Rapp attended three meetings at the Goddard Space Flight Center and one meeting at the Defense Mapping Agency Aerospace Center. Professor Jekeli attended two meeting at the Goddard Space Flight Center.

5. Publications and Presentations

The following list summarizes the publications and presentations that were made describing activities carried out under this project:

Balasubramania, N., Definition and Realization of a Global Vertical Datum, Phillips Laboratory Scientific Report No. 1, PL-TR-94-2144, April 1994; also Report No. 427, Department of Geodetic Science and Surveying, The Ohio State University, Columbus, 117p., April 1994, ADA283303.

Jekeli, C., Harmonic analysis, aliasing, and the mean anomaly. Presented at the meeting of Working Group I, Combination Methods for High-Degree Expansions, NASA/DMA Joint Gravity Field and Geoid Improvement Project, 4 October 1994, NASA,GSFC, 1994.

Jekeli, C., Spherical harmonic analysis, aliasing, and filtering. In press, *Manuscripta Geodaetica.*, 1995

Rapp, R.H., A World Vertical Datum Proposal, prepared for IAG Special Study Group 5.149, IOS Deacon Laboratory, Wormley, England, December 1993

Rapp, R.H., A Global Vertical Reference Frame, prepared for Geodetic Fixing of Tide Gauge Bench Marks, meeting, Wormley, England, December 1993

Rapp, R.H., Vertical Datums in Land and Marine Positioning, in Prod. of International Symposium on Marine Positioning, Hannover, p. 95-99, Marine Technology Society, 1994

Rapp, R.H. with R.S. Nerem, A Joint GSFC/DMA Project for Improving the Model of the Earth's Gravitational Field, presented at the meeting of the International Geoid Commission, Graz, Austria, September 1994

Rapp, R.H., The Use of Potential Coefficient Models in Computing Geoid Undulations, prepared, in part, for the International School for the Determination and Use of the Geoid, October, 1994; to appear in the Bulletin Series of the International Geoid Service, 1995

Appendix

This appendix contains the following material prepared under this contract:

Rapp, R.H.,	A World Vertical Datum Proposal, December 1993
Rapp, R.H.,	Vertical Datums in Land and Marine Positioning, September 1994.
Jekeli, C.,	Harmonic Analysis, Aliasing, and the Mean Anomaly, October 1994.
Jekeli, C.,	Spherical Harmonic Analysis, Aliasing, and Filtering, 1995.

A World Vertical Datum Proposal

Richard H. Rapp
Department of Geodetic Science & Surveying
The Ohio State University
Columbus, Ohio 43210

prepared for presentation at the meeting of IAG/Special Study Group 5.149 - Vertical Datum Investigation
IOS Deacon Laboratory
Wormley, Godalming, England

December 17, 1993

Aspects of Vertical Datum

- Origin of Vertical Datums Associated with Mean Sea Level
- Mean Sea Level Does Not Form an Equipotential Surface
- Regional Vertical Datums Developed by Fixing One or More Heights (Potential) at Tide Gauge Stations
- Several Hundred Vertical Datums in the World

Vertical Datums in Practice

- Country or Area Datums

NAVD88	North America
NAGD29	North America
AHD71	Australia
IGN69	France
ODN (Newlyn)	England
NN	Germany
Dansk Normal Nul	Denmark
N60	Finland
RH70	Sweden

- Scientific Datum

UELN

Vertical Datum Problems

- Regional Datums Have Different Reference Surfaces
- No Easy Connection for Vertical Datums
- No Consistent Technique for Datum Determination (e.g. one or several tide gauge heights held final.)
- Datums Tied to Mean Sea Level which:

  A. Changes in Time

  B. Is Not Available in All Areas

Regional Datum Connections

- Let h = ellipsoidal height
  N = geoid undulation
  H_D = orthometric height with respect to a specific datum

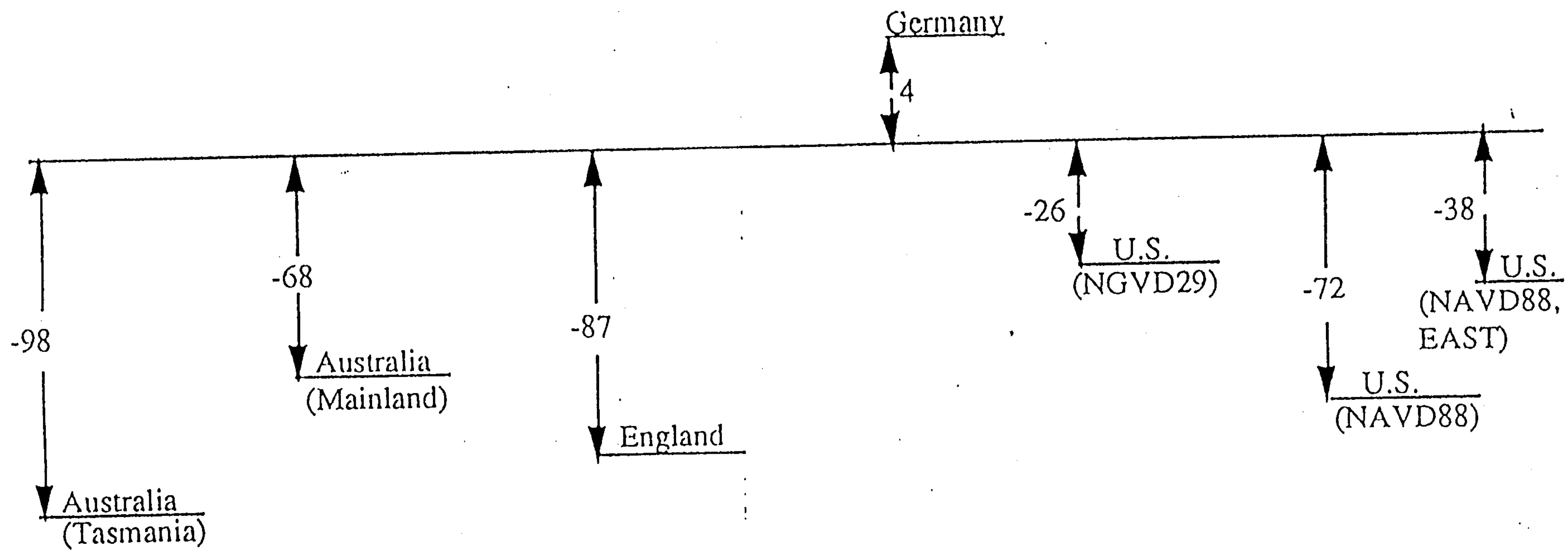
- Let B be the separation between the reference surface of the regional datum and the ideal datum

  $$H = H_D + B$$

- Given h values on related datums, values of N, B values can be determined.

Data Used for Tests

- Approximately 2000 Doppler positioned stations from 1971-1986
- Stations positions corrected for scale and translation effects to go into ITRF90 and height computed with respect to an ellipsoid with a = 6378136.3 m
- N values determined from merger of JGM-2 and OSU91A potential coefficient model which is complete to degree 360
- Data deleted if residuals greater than 2 m
- RMS Differences: ±99 cm (1417 stations)



LOCATION OF REGIONAL VERTICAL DATUM REFERENCE SURFACE WITH RESPECT TO THE GEOID

(Units are cm)

Selected Connection Results

- Germany/England Datum Difference

91 cm	Doppler/Gravity
80 cm	Leveling

(German datum above England)

- Australia (AHD71) with respect to Geoid

- 68 cm	Doppler/Gravity
106 cm	Laser/Gravity

Ideal World Height System (WHS)

- For many years proposals have been made that would lead to the definition of a WHS

- Data to be Used

  Geocentric Positions
  Mean Sea Level (Tide Gauges)
  Gravity Field Information (for Geoid Undulation)
  Satellite Altimeter Data
  Dynamic Ocean Topography

- Difficulty in acquiring well distributed and sufficiently accurate data

A Simplified World Height System

- h = ellipsoidal height
  H = orthometric height
  H^* = normal height
  N = geoid undulation
  ζ = height anomaly

  $$h = H + N = H^* + \zeta$$

  h is referred to a defined ellipsoid, preferably the "best" ellipsoid

- Given h and N calculate H

  $$H = h - N$$

- Given h and ζ calculate H^*

  $$H^* = h - \zeta$$

What is the Reference Surface?

- The proposed WHS introduces the geoid as the reference equipotential surface.
- The geoid is not directly tied to mean sea level. It is indirectly tied to an ideal reference surface by the proper choice of the ellipsoid parameters.
- The equatorial radius can be determined to an accuracy of about ±15 cm, perhaps better.

Accuracy of Geoid Determination - 1993

- Absolute Accuracy vs Relative Accuracy
- Absolute Accuracy Needed at Some Sites
- Relative Accuracy Needed for Densification
- Absolute Accuracy Today From Global 360 Models

  Land: $\pm$40 cm to $\pm$5 m
  Ocean: $\pm$30 cm

- Incorporation of Local Gravity Data Improves Accuracy

Accuracy of Geoid Determination - 1995

- If a substantial effort is made to improve the degree 360 representation, we can expect an improvement in our geoid undulation accuracy especially in areas lacking terrestrial gravity data in existing (1993) models.

- Anticipated Estimates - 360 Model

  Land: ±30 cm to ±100 cm
  Ocean: ±25 cm

- Improved accuracy with detailed gravity data

World Height System 1995(?)

- This WHS will be based on satellite positioning (primarily GPS) and geoid undulation determinations from a degree 360 model from DMA/NASA.

- The geoid undulations will be originally determined in the best reference frame and best system of constants. Grids will be created in a system compatible with the frame and constants of WGS84.

- Transformation functions can be developed to convert the WHS95 system to local datums or vice versa when needed.

Conclusions

- Progress towards an Ideal WHS has been slow.
- A global WHS is needed to add a "mean sea level" height component to WGS84.
- The geoid is a logical choice for the fundamental reference surface although it is not physically realizable.
- To achieve the accuracy goals for WHS95, a significant effort is needed to improve the representation of the Earth's gravitational potential. Such an effort is now being planned as a joint NASA/DMA project.

Vertical Datums in Land and Marine Positioning[1]

Richard H. Rapp
Department of Geodetic Science and Surveying
The Ohio State University
Columbus, Ohio 43210

ABSTRACT

Vertical datums provide the reference surface for measuring heights of the topography. The heights are usually expressed as mean sea level heights or orthometric heights or normal heights, etc. The reference surface for the vertical datum is ideally an equipotential surface, the geoid, of the Earth's gravity field. Since sea level (or mean sea level) deviates from the geoid by 1 to 2 meters, the various vertical datums of the world have different reference surfaces for their origin. This presentation will first consider the height differences between the reference surfaces of some selected (Australia, United States, England, Germany) vertical datums. The discussion will then turn to the need for a unique vertical datum in land areas and then ocean regions. Whether one considers land or ocean applications, a case will be made that the geoid should be adopted as the reference surface recognizing that the absolute portions of the geoid can only be determined to an absolute accuracy of 10 to 100 cm depending on data available for use. To achieve such accuracy at all wavelengths, improvements in our knowledge of the Earth's gravity field are needed. Some improvements can be made processing existing or soon to be available data.

Introduction

Heights and depths are a common quantity of vast usefulness. Historically height determination was referenced to mean sea level which was accessible through tide gauge measurements. Heights were then given with respect to mean sea level. Although one finds normal heights and orthometric heights in use today, they have a common need for a suitable reference surface. As positioning needs evolved numerous vertical datums were developed for various countries, regions, states within countries, etc. Today there are 100-200 different vertical datums in the world.

The concept of the datum is to have a gravity equipotential surface from which orthometric heights or geopotential numbers can be measured. Since mean sea level lies on different equipotential surface at different locations, those datums based on mean sea level are not based on the same equipotential surface. Consequently, the heights from different datums are inconsistent at the ±2 meter level.

In addition, one must recognize there are some countries that do not have access to an ocean to determine a mean sea level. Such areas thus must bring to their region level lines from other areas having such access.

Interest in studying the relative locations of vertical datums and the development of a global vertical datum has been around for many years (Mather, Rizos, Morrison, 1978; Colombo, 1980; Rapp, 1983; Heck and Rummel, 1990). More accurate space positioning and improved knowledge of the Earth's gravity field has made it possible to study the differences in some major

[1]presented at the International Symposium on Marine Positioning, Hannover, Germany, September 1994

vertical datums. Two recent studies are those of Rapp (1994) and Balasubramania (1994). The relationship between the French and British vertical datums has been discussed by Willis et al. (1989). The vertical datums in the Baltic region have been studied by Kakkuri (1994) and Pan and Sjoberg (1994). The Rapp study used Doppler positioned stations while the Balasubramania study used stations more accurately positioned through satellite laser measurements. An example of results is shown in Figure 1 which corresponds to Figure A1 of the Balasubramania report.

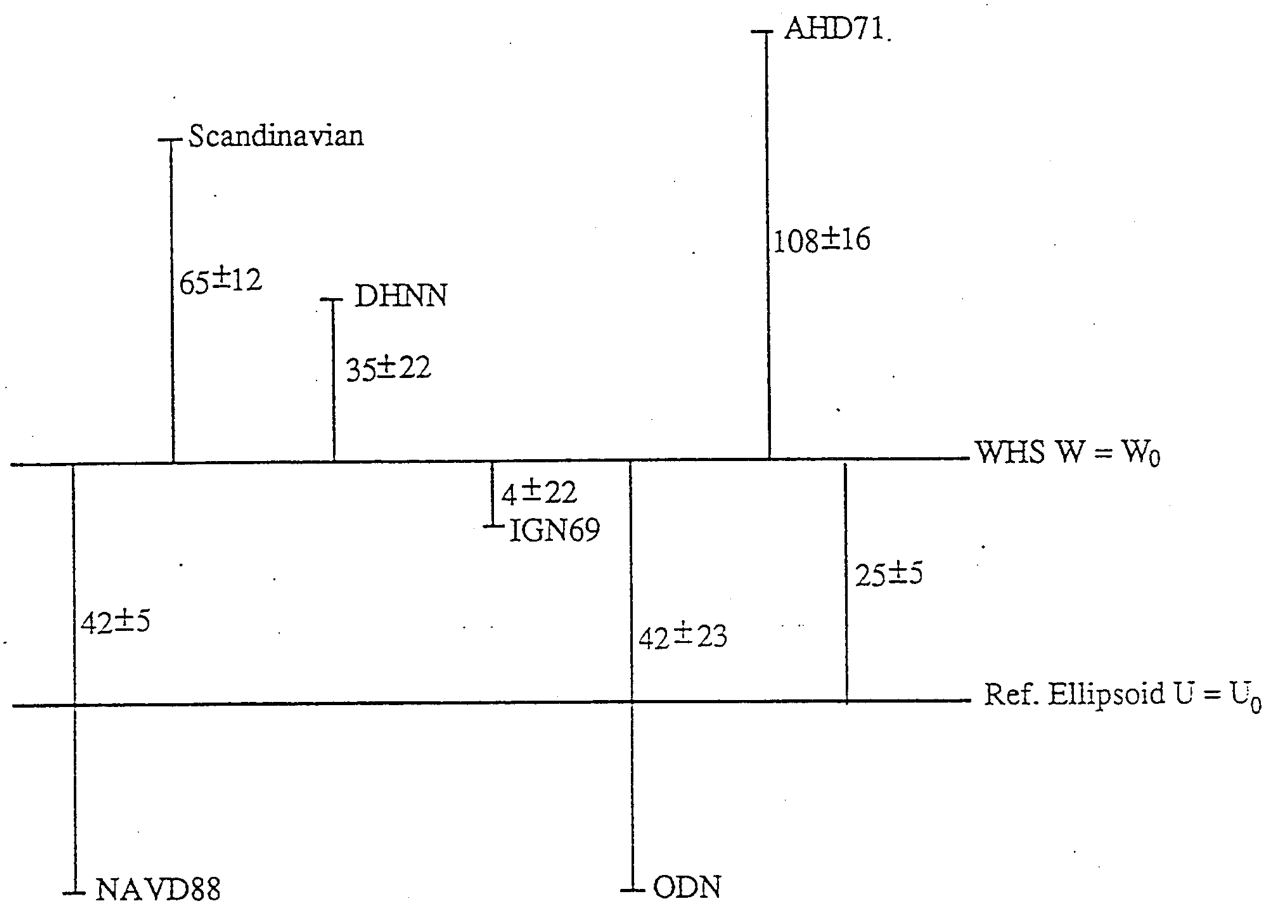


Figure 1. Vertical Datum Relationship Based on Space Located Stations. Units are cm.

This figure shows the differences between selected vertical datums and the ideal reference surface. Also shown is the reference ellipsoid (a = 6378136.3 m) adopted for these computations. One sees, for example, the origin difference between NAVD88 and the Australian datums (AHD) is 150 cm. The 25 cm difference between the ideal surface and the adopted reference ellipsoid implies that a better radius for the ellipsoid would have been 6378136.55 m which is consistent with what is being found with Topex altimeter data.

Although much attention in the past has focused on vertical datums for land applications analogous problems occur in the marine environment when bathymetry data and nautical charts are considered. In the ocean areas there are also numerous datums used in referencing a depth. These datums are usually associated with some specific tidal situation. A recent review is given is Kumar (1994).

Global Height Systems

As the precision of positioning improves with the use of GPS it becomes important to consider the determination and implementation of a global height system. Recent studies (e.g. Xu and Rummel, 1991; Rapp and Balasubramania, 1992) have examined ways to consider the global vertical datum problem. Two types of solutions are discussed. One combines a variety of space and terrestrial data to define an optimum system connected to mean sea level averaged over the oceans. Such data would include geocentric station positions, mean sea level defined at tide gauge stations, gravity field information, satellite altimeter data, and dynamic ocean topography. Although conceptually the merger of such data is desirable, there are numerous difficulties in terms of data coverage and data accuracy that may prevent the ideal solution from being realized.

An alternate approach is to use the geoid as the reference surface and let this geoid be defined by a set of geoid undulations derived gravimetrically. These undulations would then be given with respect to the best (as of some epoch and agreement) ellipsoid. The heights (h) above this ellipsoid would be determined from space techniques, primarily GPS. Knowing h and N the orthometric height can be determined:

$$H = h - N \tag{1}$$

If normal heights (H^*) are desired the gravimetric information could be used to calculate the height anomaly (ζ) so that:

$$H^* = h - \zeta \tag{2}$$

In these computations care must be taken to be in the proper reference system or frame. This system could be the satellite frame defined by WGS84 and the constants of the WGS84 ellipsoid. The height determination implied by (1) or (2) can also be applied to height difference determinations knowing that Δh can be very accurately determined through GPS procedures.

Similar procedures could take place in ocean areas for bathymetry mapping. In the case of the ellipsoid height, suitably corrected for tidal effects, the ship could be converted to a height with respect to the geoid. The depths measured by the ship could then be referred to the geoid (Kumar, 1994).

Geoid Undulation Determinations

The usefulness of this method of using the geoid as the datum origin resides on determining geoid undulations to a sufficient accuracy. Today geoid undulations can be calculated from potential coefficient models to degree 360 (Rapp, Wang, Pavlis, 1991) or the combination of such models with terrestrial gravity data. The accuracy of such determinations varies from ±30 cm in ocean areas to ±5 m in areas lacking terrestrial gravity data. This accuracy could be improved if better satellite and terrestrial data is used in the development of degree 360 models. Considering current results one could expect geoid undulations, from a new degree 360 to achieve an accuracy from ±25 cm in ocean areas to ±1.5 m in land areas not adequately covered with gravity data.

The Joint Project

Recognizing the need for precise geoid undulation determinations for both geodetic and oceanographic applications, the Defense Mapping Agency and NASA/Goddard Space Flight

Center have started a project to determine a significantly improved Earth gravity model based on a spherical harmonic representation to degree 360. The new model will use terrestrial gravity data from areas not previously available and from revised and updated files in regions currently represented in existing models. New satellite tracking data, including GPS tracking, will be incorporated in the new model. Satellite altimeter data from Topex/Poseidon, Geosat, and ERS-1 will also be used.

Activities on the development of the new model are now underway. Test solutions will be started in October 1994 with final data sets expected in July 1995. Candidate final models should be available in October 1995 with selection of the final model by early 1996. A description of the techniques and data to be used is given by Rapp and Nerem (1994).

The output of the joint project will include the potential coefficient model and a gridded geoid undulation files based only on the degree 360 model. These gridded undulations can be used for the determination of orthometric heights (or height differences) from ellipsoidal heights derived from GPS. Care will be taken to have the geoid undulations in the WGS84 system.

Conclusions

A practical implementation of a world height system or vertical datum is being considered using the geoid as the conceptual reference surface. Although this surface will not be directly tied to mean sea level, it will be associated with a global mean sea level through the proper choice of the parameters of the ideal ellipsoid. The key element in the concept will be in determining geoid undulations to meet users needs. Initially this will be done by developing a new degree 360 potential coefficient model based on significantly new data sets. This will enable undulations to be derived for most regions to an accuracy of ±1.5 m or better. Improved determinations can subsequently take place for specific regions through detailed geoid undulations computations incorporating high resolution terrestrial gravity data.

The geoid undulations will play a key role in converting ellipsoidal heights into orthometric heights with respect to the single, world wide defined, reference surface, the geoid. This will be important for both land applications and marine charting purposes. For the first time, one has the prospect of having a single vertical datum. The use of this single datum by all countries would not be realistic. However, the development of transformations between a local datum and the global datum is quite possible. This would enable the conversion of heights from a local vertical datum to the global datum or vice versa.

The increased accuracy of positioning, both on land and in the oceans, makes the development of a world height system a natural progression from a determination of a world geodetic system (e.g. WGS84). Steps are now being taken to have information to implement the world height system/vertical datum concept in 1996.

References

Balasubramania, N. (1994). Definition and Realization of a Global Vertical Datum, Report No. 427, Dept. of Geodetic Science and Surveying, The Ohio State University, Columbus, PL-TR-94-2144,ADA283303

Colombo, O.L. (1980). A World Vertical Network, Report No. 296, Dept. of Geodetic Science and Surveying, The Ohio State University, Columbus.

Heck, B. and R. Rummel (1990). Strategies for Solving the Vertical Datum Problem Using Terrestrial and Satellite Geodetic Data, in Sea Surface Topography and the Geoid, Int. Assoc. of Geodesy Symp. 104, Springer-Verlag.

Kakkuri, J. (1994). On Unification of the Vertical Datums of the Countries on the Baltic Sea, in Report 94:2, Finnish Geodetic Institute, Helsinki.

Kumar, M. (1994). A New Concept for Defining and Surveying Time-Invariant Bathymetry, 13th UN Regional Cartographic Conference for Asia and the Pacific, Beijing, May.

Mather, R., C. Rizos, and T. Morrison (1978). On the Unification of Geodetic Levelling Datums Using Satellite Altimetry, NASA Technical Memorandum 79533, Goddard Space Flight Center, Greenbelt.

Pan, M. and L.E. Sjoberg (1994). Unification of the National Height Systems with GPS - The Baltic Sea Level 1990 GPS Campaign, in Report 94:2, Finnish Geodetic Institute, Helsinki.

Rapp, R.H. (1983). The Need and Prospects for a World Vertical Datum, presented at the Symposium on the Future of Terrestrial and Space Methods for Positioning, XVII General Assembly, IAG/IUGG, Hamburg.

Rapp, R.H. (1994). Separation Between Reference Surfaces of Selected Vertical Datums, Bulletin Geodesique, (in press).

Rapp, R.H. and R.S. Nerem (1994). A Joint GSFC/DMA Project for Improving the Model of the Earth's Gravitational Field, presentation at the Int. Geoid Commission mtg., Graz, Austria, September.

Rapp, R.H., Y.M. Wang, and N.K. Pavlis (1991). The Ohio State 1991 Geopotential and Sea Surface Topography Harmonic Coefficient Models, Report No. 410, Dept. of Geodetic Science and Surveying, The Ohio State University, Columbus.

Willis, P. et al. (1989). Connection of the two Levelling System Datums IGN69 and ODN Through the Channel by Using GPS and Other Techniques, presented at the First International Workshop on Geodesy for the Europe-Africa Fixed Link Feasibility in the Strait of Gilbralter, Madrid.

HARMONIC ANALYSIS, ALIASING, AND THE MEAN ANOMALY

Issues:
- Models of Harmonic Analysis vs. Aliasing
- Models of the Mean Gravity Anomaly vs. Aliasing

Difficulty: *the sphere is a wicked surface*

- Oscar Colombo, 1980

Christopher Jekeli

Department of Geodetic Science and Surveying
The Ohio State University
Columbus OH, 43210
e-mail: cjekeli@magnus.acs.ohio-state.edu

4 October 1994

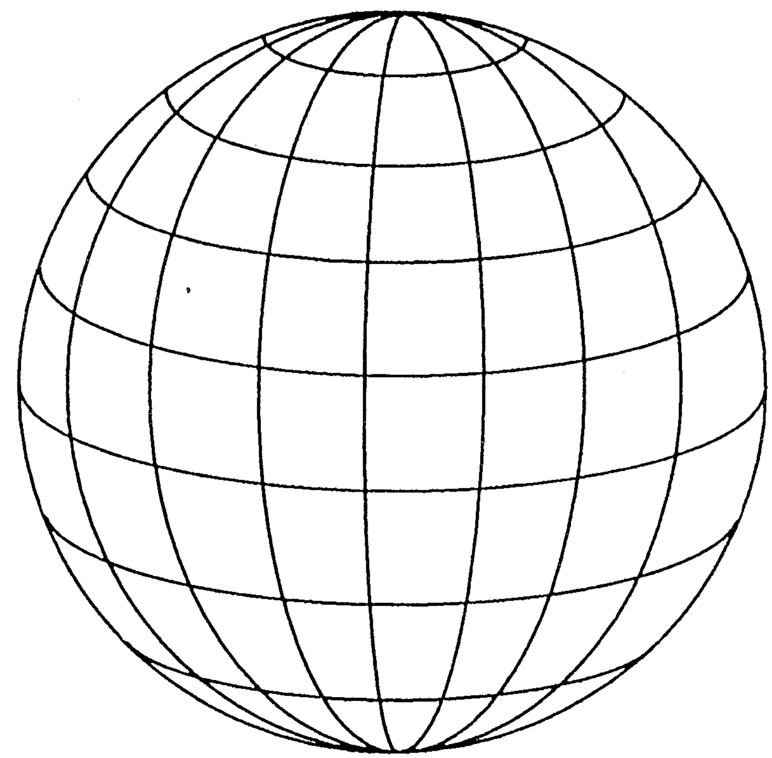
Spherical Harmonic Analysis of Gravity Anomalies

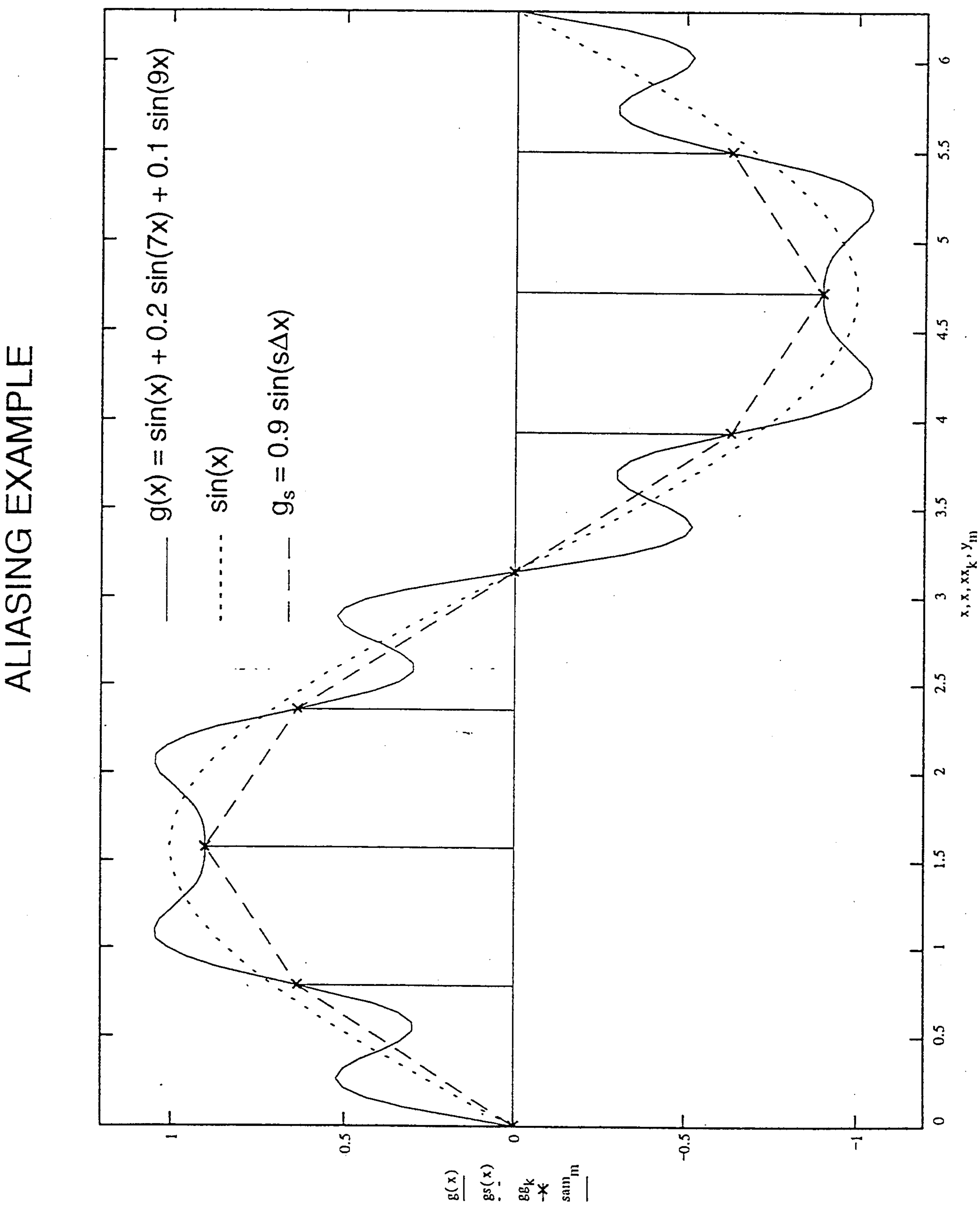
- Model: $\Delta g(\theta,\lambda) = \sum_{n=2}^{\infty} \sum_{m=-n}^{n} \gamma_{n,m}\, Y_{n,m}(\theta,\lambda)$

- Assume Regular Grid of Mean Δg Values:
  $\Delta\theta = \pi/K$, $\Delta\lambda = 2\pi/M$

- Finite Sampling ==> Aliasing

- "Eliminate" or Reduce Aliasing by Averaging Gravity Anomaly





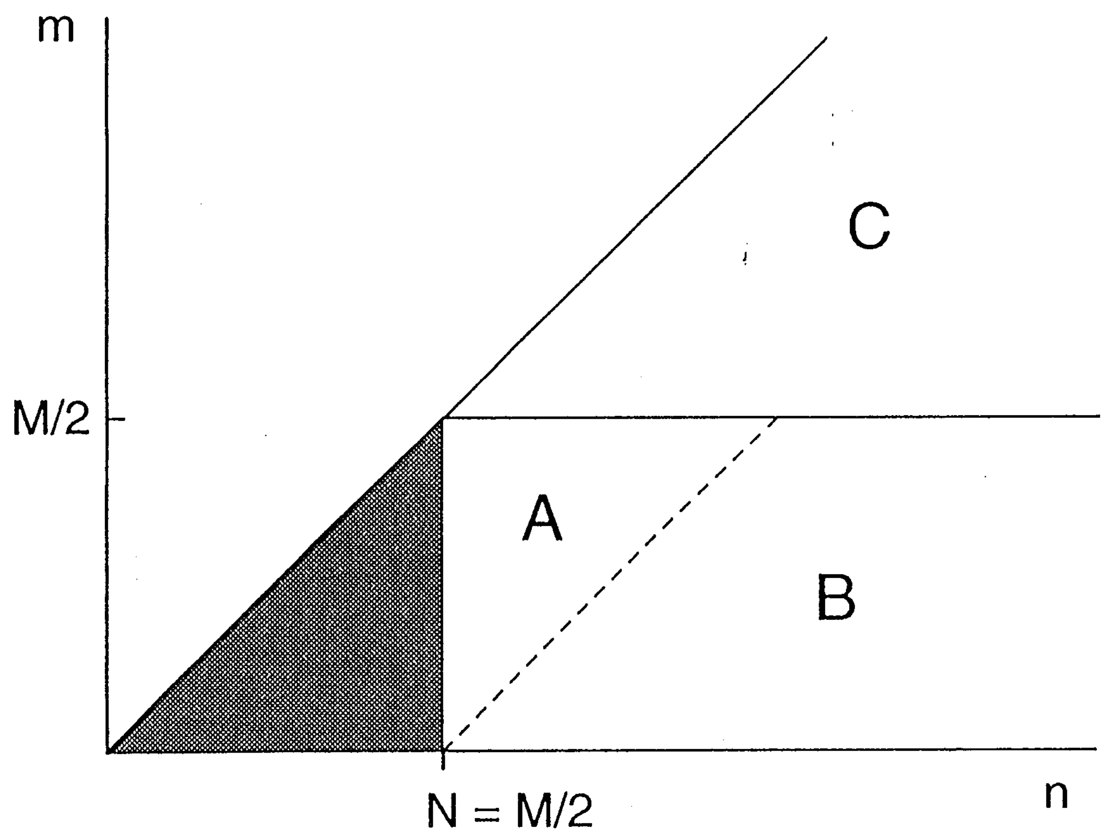
ALIASING EXAMPLE
g(x) = sin(x) + 0.2 sin(7x) + 0.1 sin(9x)
sin(x)
g_s = 0.9 sin(sΔx)
x, x, xx_k, y_m
g(x)
gs(x)
gg_k
sam_m
1
0.5
0
-0.5
-1
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
5.5
6

Models of Harmonic Analysis in Practice

Model	Deficiency
Simple Quadratures	
$\hat{\gamma}_{n,m} = \frac{1}{4\pi} \sum_{s=0}^{N-1} \sum_{t=0}^{2N-1} \Delta g(\theta_s, \lambda_t)\, Y_{n,m}(\theta_s, \lambda_t)\, \Delta\sigma_{s,t}$	Destroys Orthogonality
Conventional Least-Squares	
$\Delta g(\theta, \lambda) = \sum_{n=0}^{N} \sum_{m=-n}^{-n} \hat{\gamma}_{n,m}\, Y_{n,m}(\theta, \lambda)$	Omits High-Degree Spectrum
"New" (Extended) Model	
$\Delta g(\theta, \lambda) = \sum_{m=-N}^{N} \sum_{n=\lvert m \rvert}^{\lvert m \rvert+N} \hat{\gamma}_{n,m}\, Y_{n,m}(\theta, \lambda)$	Omits Higher-Degree Spectrum

RMS Residual Geoid Signal According to Kaula's Rule

[64/N, meters]

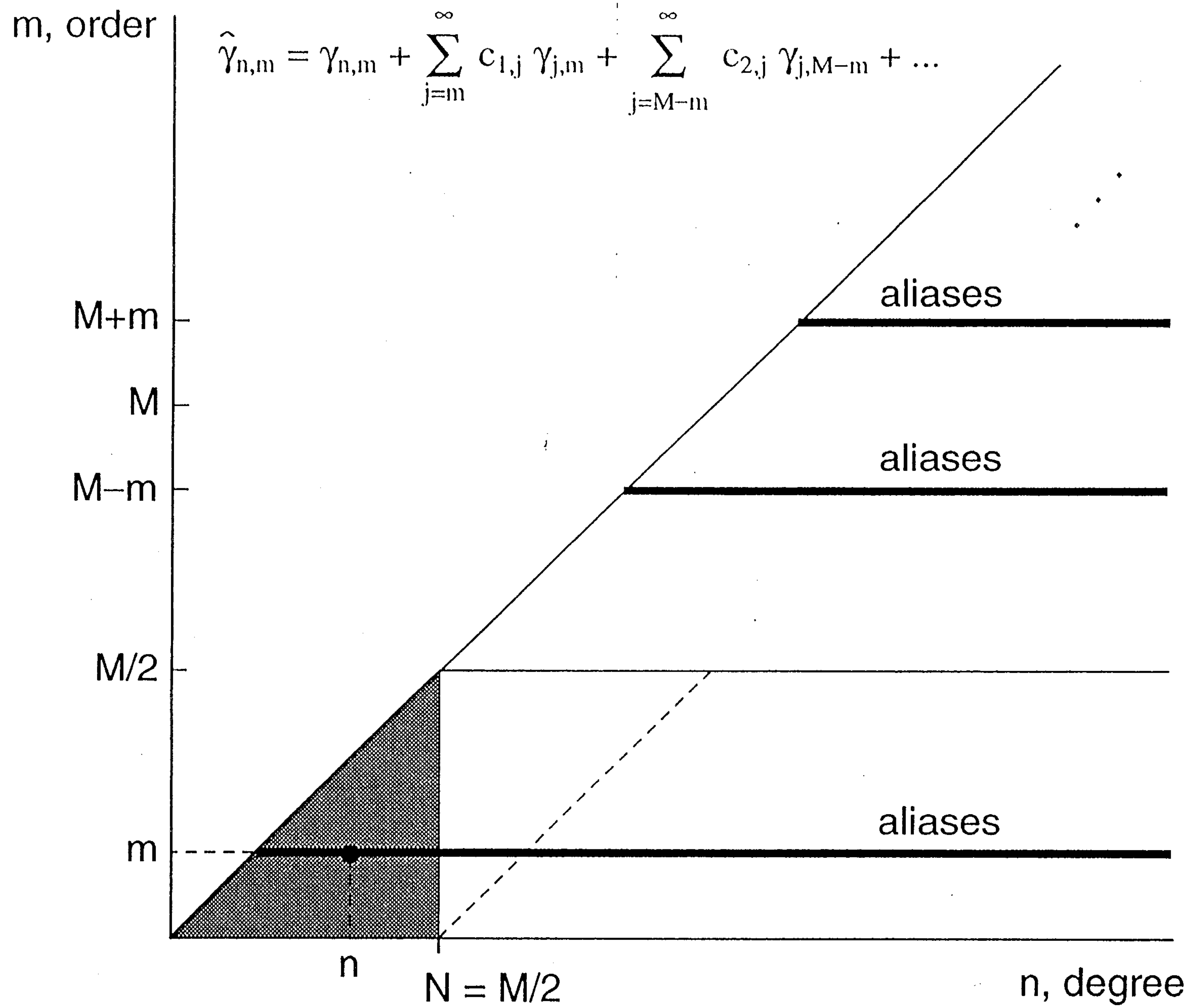


N = 360

Spectral Domain	RMS value (cm)
A + B + C	18
C	10
A + B	15
A	11
B	9
B + C	14

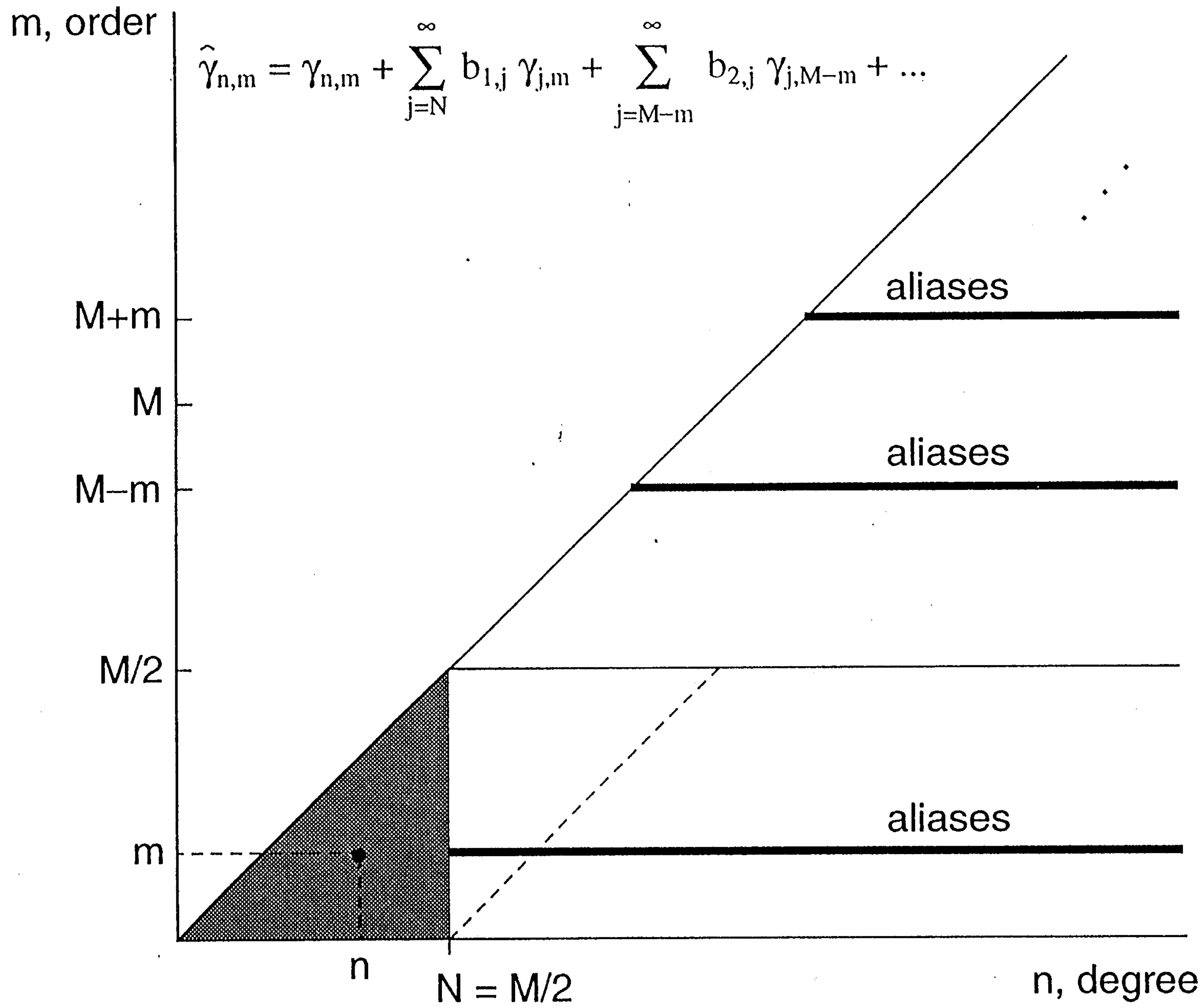
Simple Quadratures Harmonic Analysis

$$\hat{\gamma}_{n,m} = \gamma_{n,m} + \sum_{j=m}^{\infty} c_{1,j}\, \gamma_{j,m} + \sum_{j=M-m}^{\infty} c_{2,j}\, \gamma_{j,M-m} + \ldots$$



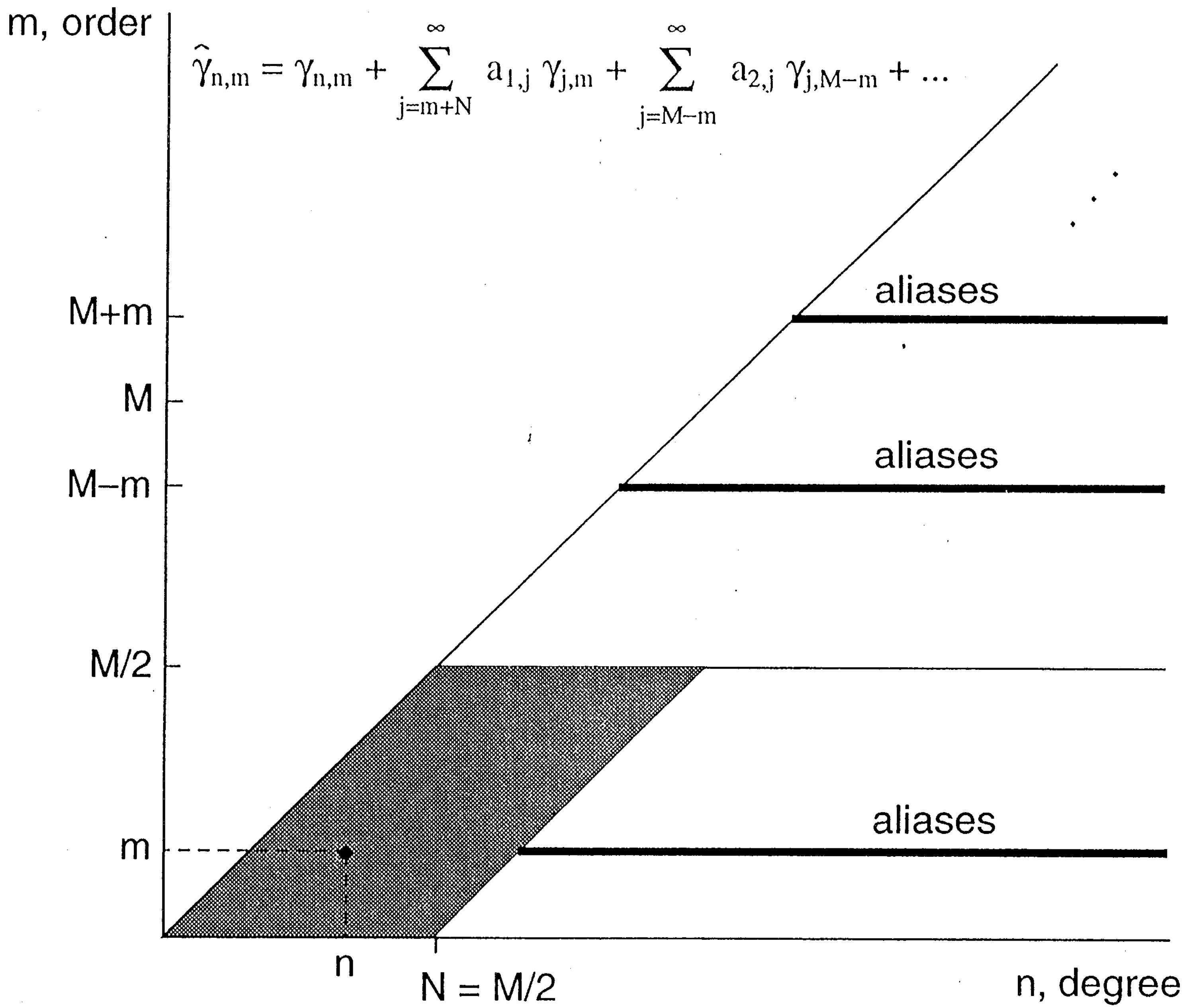
Conventional Least-Squares Harmonic Analysis

$$\hat{\gamma}_{n,m} = \gamma_{n,m} + \sum_{j=N}^{\infty} b_{1,j}\, \gamma_{j,m} + \sum_{j=M-m}^{\infty} b_{2,j}\, \gamma_{j,M-m} + \ldots$$



"New" Method of Harmonic Analysis

$$\hat{\gamma}_{n,m} = \gamma_{n,m} + \sum_{j=m+N}^{\infty} a_{1,j}\,\gamma_{j,m} + \sum_{j=M-m}^{\infty} a_{2,j}\,\gamma_{j,M-m} + \ldots$$



Options to Model the Mean Anomaly

1. Constant Angle Block Average

- Model: $\Delta\bar{g}(\theta,\lambda) = \sum_{n=2}^{\infty} \sum_{m=-n}^{n} \gamma_{n,m}\, IY_{n,m}(\theta,\lambda)$

  $IY_{n,m}$ = average of spherical harmonic function over block

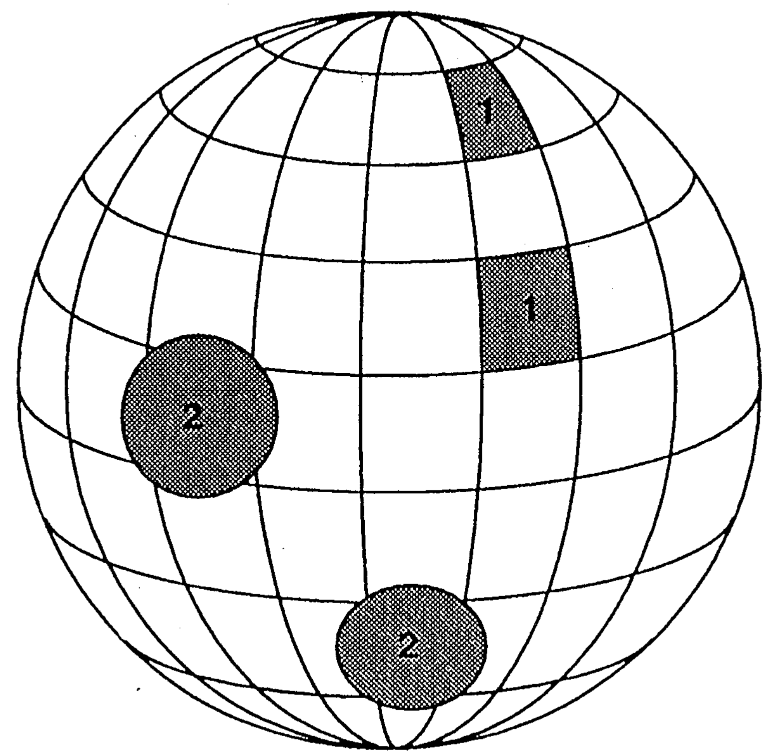
- "Eliminates" Aliases of Domain C

2. Constant Radius Cap Average

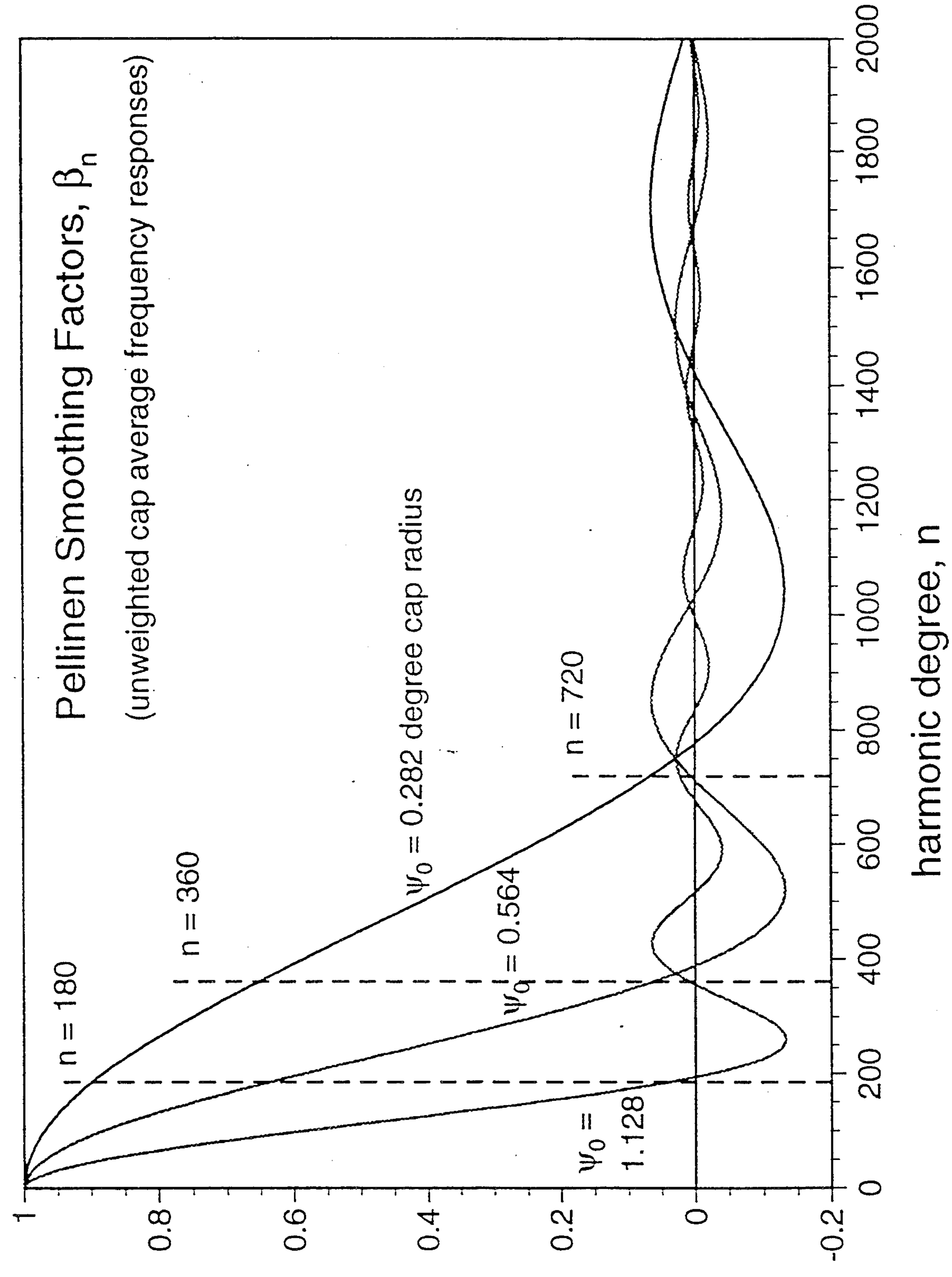
- Model: $\Delta\bar{g}(\theta,\lambda) = \sum_{n=2}^{\infty} \sum_{m=-n}^{n} \beta_n \gamma_{n,m}\, Y_{n,m}(\theta,\lambda)$

  β_n = Frequency Response of Cap Average

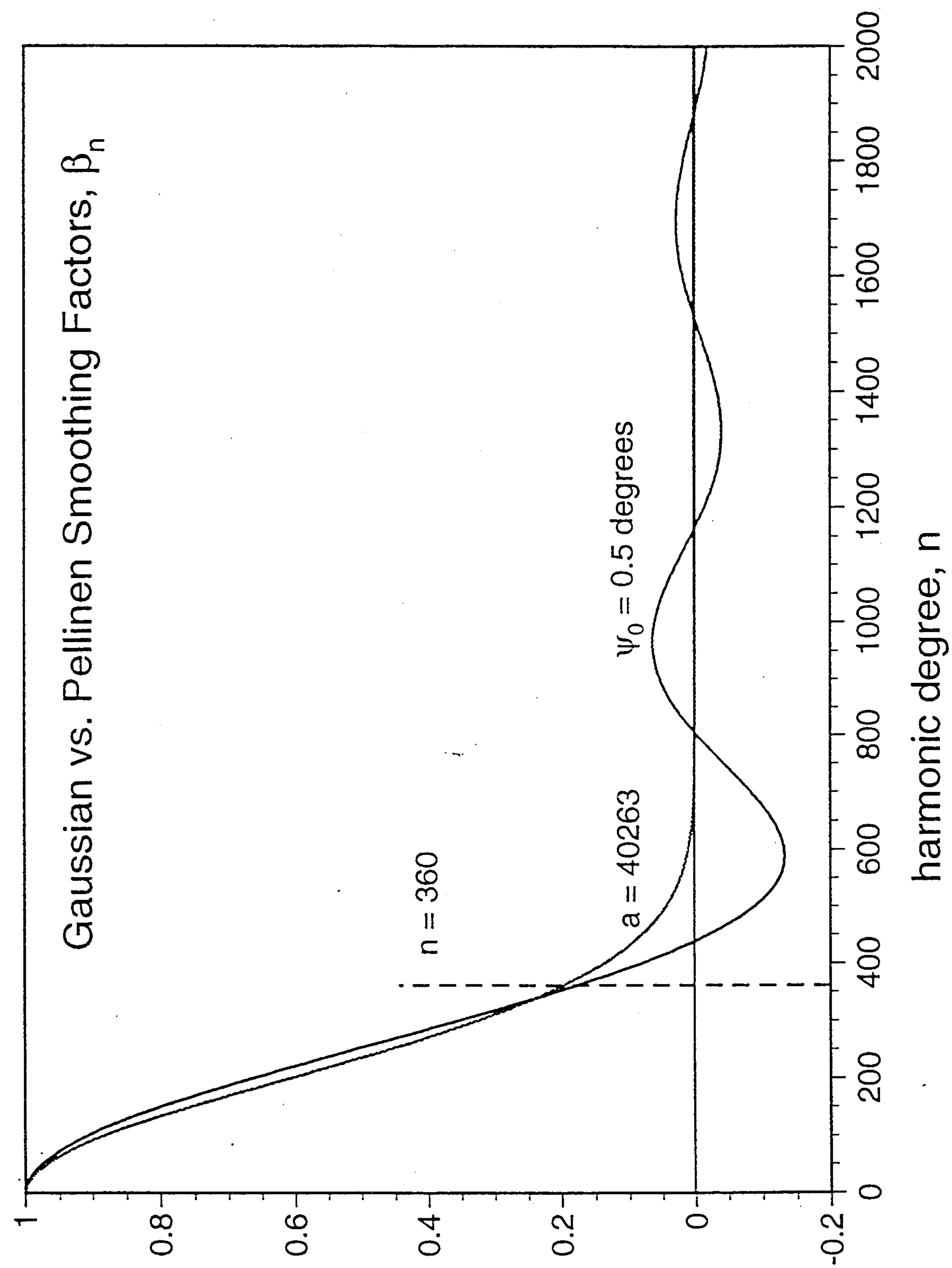
- "Eliminates" Aliases of Domain A + B + C

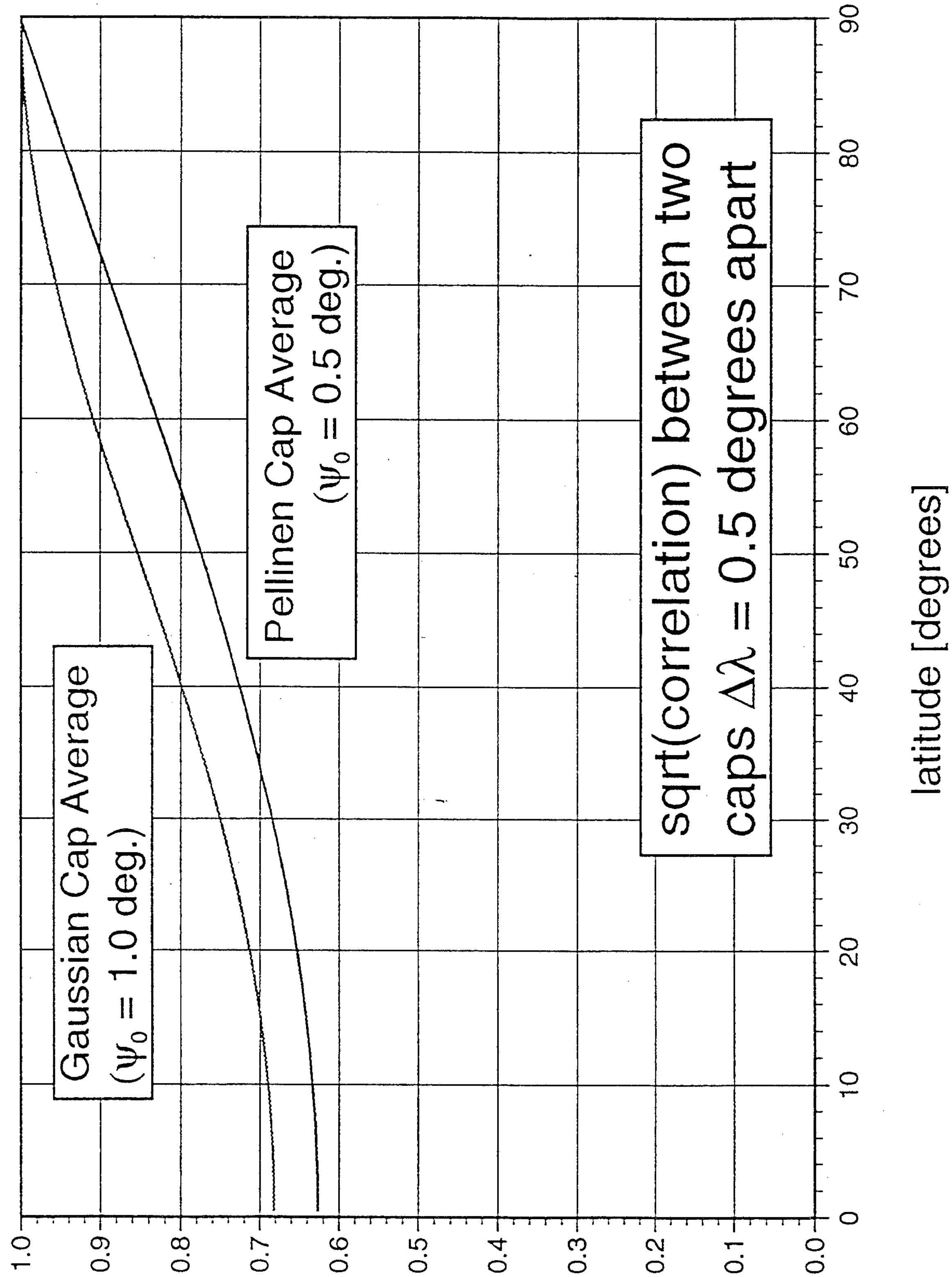


For Either Model Calculate A Regular Grid of Mean Δg Values: $\Delta\theta = \pi/K$, $\Delta\lambda = 2\pi/M$



Pellinen Smoothing Factors, β_n
(unweighted cap average frequency responses)
ψ_0 = 0.282 degree cap radius
ψ_0 = 0.564
ψ_0 = 1.128
n = 180
n = 360
n = 720
1
0.8
0.6
0.4
0.2
0
-0.2
0
200
400
600
800
1000
1200
1400
1600
1800
2000
harmonic degree, n


Gaussian vs. Pellinen Smoothing Factors, β_n
n = 360
a = 40263
ψ_0 = 0.5 degrees
harmonic degree, n
0
200
400
600
800
1000
1200
1400
1600
1800
2000
1
0.8
0.6
0.4
0.2
0
-0.2


Gaussian Cap Average
(ψ_0 = 1.0 deg.)
Pellinen Cap Average
(ψ_0 = 0.5 deg.)
sqrt(correlation) between two caps $\Delta\lambda$ = 0.5 degrees apart
latitude [degrees]
0
10
20
30
40
50
60
70
80
90
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0

Summary

- Simple quadratures (incl. variations) are deficient, in theory, for harmonic analysis
  - could be used for error propagation
- Constant angle block averages poorly attenuate the high-degree spectrum
  - may cause significant aliasing
- "New" (extended) model reduces aliasing in presence of poor filter
  - loses observational redundancy, but is a more correct model than l.s.
- Filters better than constant angle block average exist to eliminate aliasing
  - these are constant radius cap averages
  - Gaussian filter is better than Pellinen filter (unweighted average)
  - least-squares collocation is not necessarily the best

Spherical Harmonic Analysis, Aliasing, and Filtering

Christopher Jekeli
Department of Geodetic Science and Surveying
The Ohio State University
1958 Neil Ave.
Columbus, OH 43210

Abstract. The currently practiced methods of harmonic analysis on the sphere are studied with respect to aliasing and filtering. It is assumed that a function is sampled on a regular grid of latitudes and longitudes. Then, transformations to and from the Cartesian plane yield formulations of the aliasing error in terms of spherical harmonic coefficients. The following results are obtained: 1) The simple quadratures method and related methods are biased even with band-limited functions. 2) A new method that eliminates this bias is superior to Colombo's method of least squares in terms of reducing aliasing. 3) But, a simple modification of the least-squares model makes it identical to the new method as one is the dual of the other. 4) The essential elimination of aliasing can only be effected with spherical cap averages, not with the often used constant angular block averages.

Introduction

The harmonic (or spectral) representation of functions on the sphere has proved useful in modeling the gravity field, the magnetic field, the topography, the ocean tides, etc, not just for the Earth, but for any roughly spherical planet (Rapp 1977; Schmitz and Cain 1983; Kaula 1993). Such harmonic representations enable a more precise interpretation by wavelength of their magnitudes, their coherence, and their measurability. Usually, these functions are richly endowed with harmonics to a high degree, decaying relatively slowly with increasing detail. For example, the Earth's gravity field, after the first few harmonics (which certainly dominate), decays approximately linearly on the logarithmic scale, and today's representations include over 130 thousand terms (Rapp and Pavlis 1990).

It is a significant consequence of this fine structure in the functions that the harmonic components determined from a finite set of function values are biased estimates of the true components. This is a well known phenomenon in signal analysis, where even though one applies orthogonal operators to the data, they cannot filter out harmonics finer in detail than that dictated by the sampling interval. This is known as "aliasing." Another interpretation, though less precise, is that of attempting to determine all (possibly an infinite number of) harmonics from a given number of function values. In the general case, this is an underdetermined problem and the solution invariably lumps together harmonics in a prescribed way. Spectral aliasing in functions defined on the line or on the Cartesian plane is well understood; it is less obvious on a non-Euclidean surface such as the sphere.

The spherical harmonic analysis is examined here with the object of clearly understanding the effect of aliasing in conventional techniques, as well as in the modern techniques developed by Colombo (1981) that are used to compute today's models. In so doing, a new formula for the analysis is rendered that reduces the aliasing effect better than the least-squares approach. It is, however, also shown that the new technique is equivalent to a modification of the least-squares model, requiring some extra computations. The results also indicate clearly which types of averaging (filtering) do and do not eliminate aliasing. It is assumed throughtout that the data are provided discretely at every node of a uniform grid. (Aliasing in the case of a nonuniform data distribution was studied by Sansò 1990.) Though the errors in measuring the function on the sphere are not considered initially, their effect on the derived procedures is discussed in conclusion.

Preparatory Concepts

A periodic function integrable over its period may be represented either in terms of its independent (space or time) variable or, where it is continuous, in terms of its Fourier spectrum, being the set of coefficients in its corresponding Fourier series (Priestley 1981):

$$g(x) = \sum_{k=-\infty}^{\infty} G_k \, e^{i2\pi kx/T} , \qquad x \in \mathcal{R} \tag{1a}$$

$$G_k = \frac{1}{T}\int_0^T g(x)\, e^{-i2\pi kx/T}\, dx , \qquad k \in \mathcal{Z} \tag{1b}$$

where $i^2 = -1$, T is the period ($g(x+T) = g(x)$), $\mathcal{R}$ is the set of real numbers, $\mathcal{Z}$ is the set of integers, and the integer k may be termed "frequency," or "wavenumber." In practice, for example from measurements, one knows only discrete values of g and the corresponding spectrum is the Discrete Fourier Transform (DFT):

$$g_s = \sum_{k=0}^{K-1} \mathring{G}_k \, e^{i2\pi ks/K} , \qquad s = 0,\ldots,K-1 \tag{2a}$$

$$\mathring{G}_k = \frac{\Delta x}{T} \sum_{s=0}^{K-1} g_s \, e^{-i2\pi ks/K} , \qquad k = 0,\ldots,K-1 \tag{2b}$$

where it is assumed that the function values are known on a regular grid, i.e., one with constant spacing, $\Delta x = T/K$. The unsymmetric transform (2a,b), where k is nonnegative, is preferred here as being more conventional and avoiding the separate treatment of even and odd K. Since the DFT is periodic with frequency period K:

$$\mathring{G}_{k+K} = \mathring{G}_k , \qquad \text{for any } k \tag{2c}$$

it is noted that equations (2a,b) are equivalent to the symmetric forms, where k (and s) ranges from $-(K-1)/2$ to $(K-1)/2$ (if K is odd). In the notation of (2a,b), the frequencies from $(K+1)/2$ to $K-1$ are used to represent the negative frequencies from $-(K-1)/2$ to -1.

The relationship between the spectra (1b) and (2b) can be found by substituting $x = s\Delta x$ into (1a):

$$\begin{aligned} g_s \equiv g(s\Delta x) &= \sum_{j=-\infty}^{\infty} G_j \, e^{i2\pi js/K} \\ &= \sum_{k=0}^{K-1} \sum_{j=-\infty}^{\infty} G_{jK+k} \, e^{i2\pi ks/K} \end{aligned} \tag{3}$$

from which, by comparing with (2a), one has for any k,

$$\mathring{G}_k = \sum_{j=-\infty}^{\infty} G_{jK+k} = G_k + \sum_{\substack{j=-\infty \\ j\neq 0}}^{\infty} G_{jK+k} \tag{4}$$

The infinite sum on the far right in (4) is an *alias* of G_k for any frequency, k, such that $|k| < k_N = K/2$, where k_N is the so-called Nyquist frequency. The determination of the spectrum $\{G_k\}$ of the function g from the discrete sequence $\{g_s\}$, using (2b), is subject to an *aliasing error* as formulated in (4). Note that if the function g does not contain harmonics with frequency above the Nyquist frequency, then there is no aliasing error. Such functions will be called band-limited. It is further noted that a function with infinite bandwidth can be filtered to make it band-limited. This procedure is discussed later with respect to the particular application of spherical harmonic analysis.

Obviously, similar formulas hold for periodic functions defined on the Cartesian plane. They are specialized here to the case where g is a *real* function of two variables, θ and λ, and is periodic in θ with period π and in λ with period 2π. Then

$$g(\theta,\lambda) = \sum_{k=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} G_{k,m} \, e^{i(2k\theta + m\lambda)} , \qquad \theta,\lambda \in \mathcal{R} \tag{5a}$$

$$G_{k,m} = \frac{1}{2\pi^2}\int_0^{2\pi}\int_0^{\pi} g(\theta,\lambda)\, e^{-i(2k\theta + m\lambda)}\, d\theta d\lambda , \qquad k,m \in \mathcal{Z} \tag{5b}$$

where, with * denoting the complex conjugate,

$$G_{k,-m} = G^{*}_{-k,m} \tag{5c}$$

For a regular grid with constant spacings given by $\Delta\theta = \pi/K$ and $\Delta\lambda = 2\pi/M$, where $\theta_s = s\Delta\theta$, $\lambda_t = t\Delta\lambda$, the Discrete Fourier Transform pair is

$$g_{s,t} \equiv g(\theta_s,\lambda_t) = \sum_{k=0}^{K-1} \sum_{m=0}^{M-1} \mathring{G}_{k,m} \, e^{i2\pi(sk/K + tm/M)} , \qquad s = 0,\ldots,K-1 , \; t = 0,\ldots,M-1 \tag{6a}$$

$$\mathring{G}_{k,m} = \frac{1}{KM} \sum_{s=0}^{K-1} \sum_{t=0}^{M-1} g_{s,t} \, e^{-i2\pi(sk/K + tm/M)} , \qquad k = 0,\ldots,K-1 , \; m = 0,\ldots,M/2 \tag{6b}$$

Because the values $g_{s,t}$ are real,

$$\mathring{G}_{k,M-m} = \mathring{G}^{*}_{K-k,m} \tag{6c}$$

which explains the limit of interest on m. M is assumed to be even, without loss in generality. By the periodicity in k and m, $\mathring{G}_{0,m} = \mathring{G}^{*}_{0,M-m}$ and $\mathring{G}_{k,0} = \mathring{G}^{*}_{K-k,0}$, while $\mathring{G}_{0,0}$ is real. Hence, the complex, discrete spectrum, $\{\mathring{G}_{k,m}\}$, contains only KM independent coefficients. Analogous to (4), the relationship between the total spectrum (5b) and

the sample spectrum (6b) is, for any k and m,

$$\overset{\circ}{G}_{k,m} = G_{k,m} + \sum_{\substack{u=-\infty \\ u+v\neq 0}}^{\infty} \sum_{v=-\infty}^{\infty} G_{uK+k,vM+m} \tag{7}$$

The double sum in (7) represents the aliasing error for frequencies $|k| < K/2$ and $|m| < M/2$.

Instead of the plane, the real function $g(\theta,\lambda)$ can also be defined on a unit sphere and then expressed in terms of a series of spherical harmonics rather than sinusoids. There are many variations in such a representation. For simplicity, the complex exponential form is used here and its relationship to the more customary form is also given for the reader's convenience. The spherical harmonic series defines the complex Legendre spectrum $\{\gamma_{n,m}\}$ of the function as follows:

$$g(\theta,\lambda) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \gamma_{n,m} P_{n,|m|}(\cos\theta)\, e^{im\lambda} \;; \qquad 0 \le \theta \le \pi \,,\; 0 \le \lambda \le 2\pi \tag{8a}$$

$$\gamma_{n,m} = \frac{1}{4\pi\varepsilon_m} \int_0^{2\pi} \int_0^{\pi} g(\theta,\lambda) P_{n,|m|}(\cos\theta)\, e^{-im\lambda} \sin\theta \, d\theta \, d\lambda \;; \qquad n \ge 0 \,,\; -n \le m \le n \tag{8b}$$

where

$$\varepsilon_m = \begin{cases} 1 \,, & m = 0 \\ 2 \,, & m \neq 0 \end{cases} \tag{8c}$$

and, since g is real,

$$\gamma_{n,-m} = \gamma^*_{n,m} \tag{8d}$$

The integers n and m are, respectively, the degree and order of the harmonic coefficient, $\gamma_{n,m}$. The $P_{n,m}$ in (8a) an (8b) are called associated Legendre functions of the first kind, defined, for the present purpose, for non-negative integers n,m with $m \le n$ and for $0 \le \theta \le \pi$. Furthermore, they are assumed to be normalized such that

$$\frac{1}{2} \int_0^{\pi} \left(P_{n,m}(\cos\theta)\right)^2 \sin\theta \, d\theta = \varepsilon_m \tag{9}$$

(again, for simplicity, the conventional over-bar notation is omitted). The complex harmonic coefficients are related to the coefficients of Heiskanen and Moritz (1967, p.31, (1-76)) by

$$\gamma_{n,m} = \begin{cases} \frac{1}{2}\left(\bar{a}_{n,m} - i\, \bar{b}_{n,m}\right), & m > 0 \\ \bar{a}_{n,0} \,, & m = 0 \\ \frac{1}{2}\left(\bar{a}_{n,m} + i\, \bar{b}_{n,m}\right), & m < 0 \end{cases} \tag{10}$$

Because of (8d), one need only solve for coefficients with non-negative orders although, formally, the series must include $m < 0$.

The Fourier Spectra of $P_{n,m}$ and $\sin\theta\, P_{n,m}$

It will be necessary to know the Fourier spectra of $P_{n,m}$ and of $\sin\theta \cdot P_{n,m}$, both defined for $0 \le \theta \le \pi$, as usual, and for $\theta < 0$ and $\theta > \pi$ by periodic extension. This means that the number of frequencies of the spectrum in each case is *infinite*, even though these functions are polynomials of sinusoids (Hobson 1965, p.96). For $0 \le \theta \le \pi$, and any integer, k, let

$$P_{n,m}(\theta) = P_{n,m}(\cos\theta) \quad \text{and} \quad P_{n,m}(\theta + k\pi) = P_{n,m}(\theta) \tag{11a}$$

$$S_{n,m}(\theta) = \sin\theta\, P_{n,m}(\cos\theta) \text{ and } S_{n,m}(\theta + k\pi) = S_{n,m}(\theta) \tag{11b}$$

In accord with (1a) and (1b), the corresponding Fourier transform pairs are

$$P_{n,m}(\theta) = \sum_{k=-\infty}^{\infty} \rho_k^{n,m}\, e^{i2k\theta} \,, \quad \theta \in \mathcal{R} \tag{12a}$$

$$\rho_k^{n,m} = \frac{1}{\pi} \int_0^{\pi} P_{n,m}(\theta)\, e^{-i2k\theta}\, d\theta \,, \quad k \in \mathcal{Z} \tag{12b}$$

and

$$S_{n,m}(\theta) = \sum_{k=-\infty}^{\infty} \sigma_k^{n,m}\, e^{i2k\theta} \,, \quad \theta \in \mathcal{R} \tag{13a}$$

$$\sigma_k^{n,m} = \frac{1}{\pi} \int_0^{\pi} \sin\theta\, P_{n,m}(\theta)\, e^{-i2k\theta}\, d\theta \,, \quad k \in \mathcal{Z} \tag{13b}$$

It is noted that because $P_{n,m}$ and $S_{n,m}$ are real functions,

$$\rho_k^{n,m} = \left(\rho_{-k}^{n,m}\right)^* \quad \text{and} \quad \sigma_k^{n,m} = \left(\sigma_{-k}^{n,m}\right)^* \tag{14}$$

The orthogonality of the associated Legendre functions

implies that the infinite sequences $\{\rho_k^{n,m}\}$ and $\{\sigma_{-k}^{j,m}\}$ are orthogonal for $n \neq j$:

$$\sum_{k=-\infty}^{\infty} \rho_k^{j,m}\, \sigma_{-k}^{n,m} = \frac{2\varepsilon_m}{\pi}\, \delta(j,n) \tag{15}$$

where $\delta(j,n)$ is the delta function, equal to zero unless $j = n$, when it is one.

Once these coefficients are computed for $n,m = 0,1$ and for all integers, k, of interest, the remaining spectra ($n,m > 1$) for these k can be determined using well known recursion formulas for the associated Legendre function (Abramowitz and Stegun 1972).

The discrete inverse Fourier transforms of these functions sampled on a uniform grid $\{\theta_s = s\,\Delta\theta : s = 0, ..., K-1;\ \Delta\theta = \pi/K\}$ are

$$P_{n,m;s} = \sum_{k=0}^{K-1} \mathring{\rho}_k^{n,m}\, e^{i2\pi sk/K} \tag{16a}$$

$$S_{n,m;s} = \sum_{k=0}^{K-1} \mathring{\sigma}_k^{n,m}\, e^{i2\pi sk/K} \tag{16b}$$

with (see (4)):

$$\mathring{\rho}_k^{n,m} = \sum_{u=-\infty}^{\infty} \rho_{uK+k}^{n,m}\ ; \quad \mathring{\sigma}_k^{n,m} = \sum_{w=-\infty}^{\infty} \sigma_{wK+k}^{n,m} \tag{17}$$

The discrete spectra in (17) are also amenable to computation by recursion formulas (see Appendix) since the left-hand sides are simply finite sums of associated Legendre functions. Other methods to compute these include simply determining the Discrete Fourier Transform (A.1) of the Legendre functions determined from recursion formulas, or substituting their equivalent finite polynomial of sinusoids into (A.1).

The Relationships Between the Fourier and Legendre Spectra

The following orthogonality relationships hold:

$$\frac{1}{2\pi}\int_0^{2\pi} e^{i(m'-m)\lambda}\, d\lambda = \begin{cases} 0\,, & \text{if } m \neq m' \\ 1\,, & \text{if } m = m' \end{cases} \tag{18a}$$

and

$$\frac{1}{\pi}\int_0^{\pi} e^{i2(k'-k)\theta}\, d\theta = \begin{cases} 0\,, & \text{if } k \neq k' \\ 1\,, & \text{if } k = k' \end{cases} \tag{18b}$$

and for later use:

$$\sum_{t=0}^{M-1} e^{i(j-m)\lambda_t} = \begin{cases} 0\,, & j \neq m \\ M\,, & j = m \end{cases} \tag{18c}$$

Substituting (12a) and (8a) into (5b) and taking note of (18a,b), one obtains

$$G_{k,m} = \sum_{n=|m|}^{\infty} \gamma_{n,m}\, \rho_k^{n,|m|} \tag{19}$$

where it is further noted that $\gamma_{n,m} = 0$ for $|m| > n$. Similarly, substituting (13a) and (5a) into (8b) yields the inverse relationship

$$\gamma_{n,m} = \frac{\pi}{2\varepsilon_m} \sum_{k=-\infty}^{\infty} G_{k,m}\, \sigma_{-k}^{n,|m|} \tag{20}$$

These relationships are possible because of the definitions (11a,b). The alternative suggested by Ricardi and Burrows (1972) implicitly assumes periodicity of the associated Legendre functions over $(0,2\pi)$ and, though their spectral domains are then finite, in this form they cannot be used with (5a) and (5b) that embody periodicity in latitude over $(0,\pi)$. On the other hand, Goldstein (1978) devised an extension of the function g onto the larger domain, $0 \leq \theta \leq 2\pi$, $0 \leq \lambda \leq 2\pi$, thus taking advantage of the associated Legendre functions as finite polynomials of sinusoids. In the present case the infinite nature of the their spectral domains will not cause computational problems because in the end only the discrete spectra (17) are needed. Albertella and Sacerdote (1995) also developed procedures similar to the present ones.

Quadrature Formulas

Because the function (e.g., the gravity anomaly) is supposed to be given on a regular grid, all estimates of its Legendre spectrum have the form of so-called quadrature formulas. A subset of these is the discretization of the integral (8b). Numerous variations have been proposed; for example, the simplest is

$$\mathring{\gamma}_{n,m} = \frac{\pi}{2\varepsilon_m KM} \sum_{s=0}^{K-1} \sum_{t=0}^{M-1} g_{s,t}\, P_{n,m}(\cos\theta_s)\, \sin\theta_s\, e^{-i2\pi tm/M} \tag{21}$$

More common is an attempt to incorporate the fact that usually the data are averages over angular cells. In this case, the Legendre spectrum of the *mean* anomaly is sought and the spherical harmonic function values in (21) are replaced by their averages over the cells. Either way, (20) or its alternative with averaged spherical harmonics is only an approximation and leads to estimates that are biased due to aliasing not just from frequency components higher than the Nyquist limit, but from components of all degrees greater than or equal to the order. Indeed, it is well known (Pavlis 1988; Albertella et al. 1993) that a band-limited function analyzed according to (21) is not recovered exactly by the estimated spectrum. This feature is demonstrated analytically below, but only for the case

exemplified by (21). The analogy for the case of cell-averages will be evident because the integrals of the Legendre functions have spectra like (17), but multiplied by $(\sin k\Delta\theta)/(k\Delta\theta)$.

With the aid of (13a) and (17), equation (21) becomes

$$\mathring{\gamma}_{n,m} = \frac{\pi}{2\varepsilon_m} \sum_{k=0}^{K-1} \mathring{\sigma}_{-k}^{n,|m|} \mathring{G}_{k,m} \tag{22}$$

The error in this estimate is derived by substituting (7) into (22), adding and subtracting $\gamma_{n,m}$, and making use of (17), (19) and (20). Some manipulations yield

$$\mathring{\gamma}_{n,m} = \gamma_{n,m} + \frac{\pi}{2\varepsilon_m} \left(-\sum_{k=-\infty}^{\infty} \sum_{j=|m|}^{\infty} \rho_k^{j,|m|} \sigma_{-k}^{n,|m|} \gamma_{j,m} + \sum_{k=0}^{K-1} \sum_{v=-\infty}^{\infty} \sum_{j=|vM+m|}^{\infty} \mathring{\rho}_k^{j,|vM+m|} \mathring{\sigma}_{-k}^{n,|m|} \gamma_{j,vM+m} \right) \tag{23}$$

Now shift the v=0 term to the first sum to get, with (15),

$$\mathring{\gamma}_{n,m} = \gamma_{n,m} + \frac{\pi}{2\varepsilon_m} \left[\sum_{j=|m|}^{\infty} \left(\sum_{k=0}^{K-1} \mathring{\rho}_k^{j,|\cancel{vM}+m|} \mathring{\sigma}_{-k}^{n,|m|} - \frac{2\varepsilon_m}{\pi} \delta(j,n) \right) \gamma_{j,m} + \sum_{\substack{v=-\infty \\ v \neq 0}}^{\infty} \sum_{j=|vM+m|}^{\infty} \sum_{k=0}^{K-1} \mathring{\rho}_k^{j,|vM+m|} \mathring{\sigma}_{-k}^{n,|m|} \gamma_{j,vM+m} \right] \tag{24}$$

Clearly, for band-limited data, where $\gamma_{n,m} = 0$, $n \geq M/2$, the last set of sums, where $j \geq M/2$, vanishes. But, the estimate is still biased ("aliased") by the second set that contains all harmonic coefficients with degrees $j \geq |m|$, including $\gamma_{n,m}$! The analysis procedure derived next shifts the total error of the estimate to frequencies beyond the Nyquist frequency.

A New Method

Combining (7) and (19), and using (17) leads to

$$\mathring{G}_{k,m} = \sum_{v=-\infty}^{\infty} \sum_{j=|vM+m|}^{\infty} \mathring{\rho}_k^{j,|vM+m|} \gamma_{j,vM+m} \tag{25}$$

This is the true model linking the two-dimensional discrete Fourier transform of the data (6b) to the actual Legendre spectrum. It is equivalent to the model (8a) for the nontransformed data. Again, extracting the v = 0 contribution, this can be broken into three parts:

$$\mathring{G}_{k,m} = \sum_{j=|m|}^{|m|+K-1} \mathring{\rho}_k^{j,|m|} \gamma_{j,m} + \sum_{j=|m|+K}^{\infty} \mathring{\rho}_k^{j,|m|} \gamma_{j,m} + \sum_{\substack{v=-\infty \\ v \neq 0}}^{\infty} \sum_{j=|vM+m|}^{\infty} \mathring{\rho}_k^{j,|vM+m|} \gamma_{j,vM+m} \tag{26}$$

The first sum contains exactly K harmonic coefficients for any given m, and these are to be solved, given $\mathring{G}_{k,m}$. One could have separated the first two sums in (26) at some integer larger than $|m| + K$, but then the problem is underdetermined because there are only K Fourier coefficients $\mathring{G}_{k,m}$ for any given m.

Let $\mathring{G}_m$ and γ_m be finite K-vectors containing, respectively, coefficients $\mathring{G}_{k,m}$ and $\gamma_{j,m}$ as follows:

$$\mathring{G}_m^T = \left[\mathring{G}_{0,m}, \mathring{G}_{1,m}, \ldots, \mathring{G}_{K-1,m} \right] \tag{27}$$

$$\gamma_m^T = \left[\gamma_{|m|,m}, \gamma_{|m|+1,m}, \ldots, \gamma_{|m|+K-1,m} \right] \tag{28}$$

Also, define the KxK matrix

$$R_m = \begin{bmatrix} \mathring{\rho}_0^{|m|,|m|} & \cdots & \mathring{\rho}_0^{|m|+K-1,|m|} \\ \vdots & & \vdots \\ \mathring{\rho}_{K-1}^{|m|,|m|} & \cdots & \mathring{\rho}_{K-1}^{|m|+K-1,|m|} \end{bmatrix} \tag{29}$$

Then, the collection of equations (26) for a fixed m and for k = 0, ..., K-1 is

$$\mathring{G}_m = R_m \gamma_m + \delta R_m \, \delta\gamma_m + \Delta R_m \, \Delta\gamma_m \tag{30}$$

where δR_m, $\delta\gamma_m$, ΔR_m, and $\Delta\gamma_m$ are semi-infinite matrices and vectors whose elements can be inferred from (26). In particular, $\delta\gamma_m$ contains harmonic coefficients of order m and degree greater than $|m| + K - 1$; while $\Delta\gamma_m$ contains *all* coefficients with order greater than or equal to M/2. The relative domains for these coefficients are shown in Figure 1, where it is noted that K is completely independent of M.

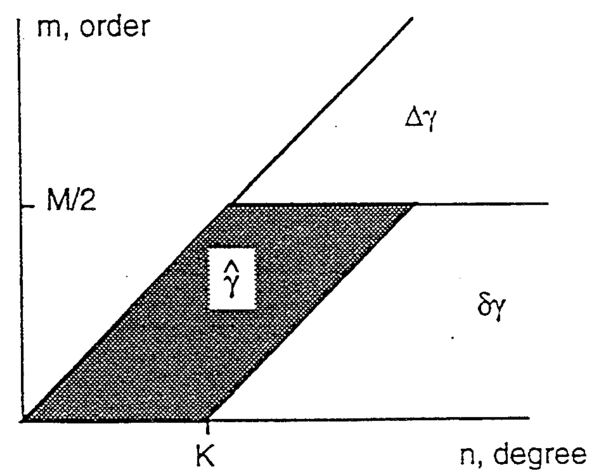


Fig.1: Domain of spherical harmonic coefficients estimated by (31). Aliasing is due to coefficients $\Delta\gamma$ and $\delta\gamma$ in their respective domains, see eq. (32).

Solving (30) for γ_m yields the following estimate of the harmonic coefficients

$$\hat{\gamma}_m = R_m^{-1} \overset{\circ}{G}_m \tag{31}$$

with an *aliasing* error given by

$$\gamma_m^a = -R_m^{-1} \left(\delta R_m \, \delta\gamma_m + \Delta R_m \, \Delta\gamma_m \right) \tag{32}$$

R_m is nonsingular as shown later (equation (49)). For a band-limited function the aliasing error (32) should vanish. Band-limited now refers to two Nyquist frequencies, just like in the planar case: $N(m) = M/2$ and $N(n) = m + K$. Thus, the spherical analysis of functions having harmonics (n,m) that are nonzero only for $n < N(n)$ and $m < N(m)$ will be perfect, without aliasing, according to (31). The use of the quadratures formula (21) would introduce errors as given by the first double sum in (24).

It is convenient, especially in efforts to eliminate the aliasing effect, to refer the Nyquist limit to a single frequency in terms of the degree, n. This is $N = \min(K, M/2)$, where usually, of course, $K=M/2$. This, however, understates the number of recoverable spectral components. That is, from the KM data, $g_{s,t}$, that determine $\overset{\circ}{G}_m$, (31) yields exactly KM real coefficients contained in the complex numbers $\hat{\gamma}_m$, $0 \le m \le M/2$; see Figure 1. Note that for each such m, the estimate (31) requires the inversion of a KxK matrix.

The Least-Squares Solution

It is now shown that the procedure above is similar to the least-squares approach developed by Colombo (1981) and implemented by Hwang (1993). However, there is a subtle difference that is easily eliminated. Equation (8a), with summations reversed, is rewritten as follows to separate symbolically the coefficients to be estimated:

$$g(\theta_s,\lambda_t) = \left(\sum_{m=-N+1}^{N-1} \sum_{n=|m|}^{N-1} + \sum_{m=-N+1}^{N-1} \sum_{n=N}^{\infty} + \sum_{|m|\ge N} \sum_{n=|m|}^{\infty} \right) \left(\gamma_{n,m} P_{n,|m|}(\cos\theta_s) e^{im\lambda_t} \right) \tag{33}$$

The usual model for the least-squares approach is the truncated spherical harmonic series, being the first part in (33):

$$g_{s,t} \approx \sum_{m=-N+1}^{N-1} \sum_{n=|m|}^{N-1} \gamma_{n,m} \, P_{n,|m|}(\cos\theta_s) \, e^{im\lambda_t} \, ; \qquad 0 \le s \le K-1, \; 0 \le t \le M-1 \tag{34}$$

The symbol $\approx$ is used to emphasize that this is not the true model, only an approximation, due to the truncation. For a more compact notation, define the following KM-vector

$$g = \left\{ g_{s,t} \right\} , \tag{35}$$

the (N - |m|)-vector (which differs from the K-vector, γ_m, by the number of elements)

$$\tilde{\gamma}_m^T = \left[\gamma_{|m|,m}, \gamma_{|m|+1,m}, \ldots, \gamma_{N-1,m} \right] , \tag{36}$$

and the KMx(N-|m|) matrix

$$\tilde{Y}_m = \left\{ P_{j,|m|}(\cos\theta_s) \, e^{im\lambda_t} \right\}_{s=0,\ldots,K-1;\; t=0,\ldots,M-1}^{j=|m|,\ldots,N-1} \tag{37}$$

The tildes (~) are included now to differentiate between this and the following modification to the approach. With (35), (36), and (37), the model (34) becomes

$$g \approx \sum_{m=-N+1}^{N-1} \tilde{Y}_m \tilde{\gamma}_m \tag{38}$$

Note that there are KM data (observations), but only approximately N^2 coefficients (if $K=M/2$, then $N^2 = KM/2$). The true model (33) can be written like (38) as follows:

$$g = \sum_{m=-N+1}^{N-1} \left(\tilde{Y}_m \tilde{\gamma}_m + \delta\tilde{Y}_m \, \delta\tilde{\gamma}_m \right) + \sum_{|m|\ge N} \Delta Y_m \, \Delta\gamma_m \tag{39}$$

where $\delta\tilde{Y}_m$ is a semi-infinite matrix with KM rows and its elements are spherical harmonic functions with degrees $n \ge N$. $\delta\tilde{\gamma}_m$ contains coefficients with degree $n \ge N$, $\Delta\gamma_m$ is the same as in (30) and ΔY_m can be inferred from (33).

By the discrete orthogonality of the complex exponential (18c), the matrices $\tilde{Y}_m$ are also orthogonal:

$$\left(\overset{\smile}{\tilde{Y}_m} \right)^T \tilde{Y}_j = 0 \, , \quad j \ne m \tag{40}$$

Multiplying (38) by $\left(\overset{\smile}{\tilde{Y}_m} \right)^T$ and denoting the normal matrix by $\tilde{A}_m = \left(\overset{\smile}{\tilde{Y}_m} \right)^T \tilde{Y}_m$ yields the "least-squares" solution:

$$\hat{\tilde{\gamma}}_m = \tilde{A}_m^{-1} \left(\overset{\smile}{\tilde{Y}_m} \right)^T g \tag{41}$$

The aliasing error is obtained by substituting the true model (39) and duly noting the orthogonalities:

$$\hat{\tilde{\gamma}}_m = \tilde{\gamma}_m + \tilde{A}_m^{-1} \left(\delta\tilde{Y}_m \, \delta\tilde{\gamma}_m + \Delta Y_m \, \Delta\gamma_m \right) \tag{42}$$

Figure 2 shows the parts of the spectrum, $\delta\tilde{\gamma}_m$ and $\Delta\gamma_m$, that alias the estimated part ($K=M/2$ is assumed). Note

the similarities of the equation pairs (41), (42) and (31), (32); however, the solution (41) is aliased more than the solution (31). Because the $\tilde{A}_m$ are (N-|m|)x(N-|m|) matrices, there are fewer computations associated with (41) (matrix inverses of diminishing dimension as m → N-1).

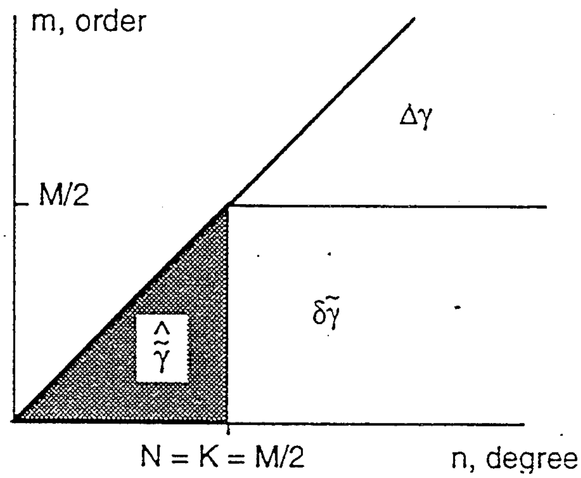


Fig.2: Domain of spherical harmonic coefficients estimated by (41), the conventional least-squares approach. Aliasing is due to coefficients $\Delta\gamma$ and $\delta\tilde{\gamma}$ in their respective domains, see eq. (42).

A Modification To The Least-Squares Model

By now it is evident that the difference between (41) and (31) lies in the model. Augmenting (34) to include degrees up to |m| + K - 1, the model becomes

$$g \approx \sum_{m=-M/2+1}^{M/2} Y_m \gamma_m \tag{43}$$

where Y_m is similar to $\tilde{Y}_m$ but now each has K columns instead of (N - |m|), and γ_m is given by (28). Correspondingly, $\tilde{A}_m$ becomes the KxK matrix A_m; and the least-squares solution (41) becomes

$$\hat{\gamma}_m = A_m^{-1} \left(Y_m^*\right)^T g \tag{44}$$

Since as many coefficients are computed as data given, this is no longer a "least-squares" solution. If g is band-limited, then both (31) and (44) yield the true spectrum, which means that (31) and (44) are identical for any g. Hence, in general, the aliasing error in this modified approach is also given by (32).

The 2-D Fourier transform $\mathring{G}_m$ (K-vector) is a linear operator on g:

$$\mathring{G}_m = \frac{1}{KM} W_m^T g \tag{45}$$

where W_m is a KMxK matrix:

$$W_m = \left\{e^{-i2\pi(sk/K + tm/M)}\right\}_{s=0,\dots,K-1;\ t=0,\dots,M-1}^{k=0,\dots,K-1} \tag{46}$$

Substituting (45) into (31) and equating with (44) yields

$$A_m^{-1} \left(Y_m^*\right)^T g = \frac{1}{KM} R_m^{-1} W_m^T g \tag{47}$$

which holds for arbitrary g. Therefore, since $W_m \left(W_m^*\right)^T = KM\ I$, where I is the identity matrix,

$$A_m = \left(Y_m^*\right)^T \left(W_m^*\right)^T R_m \tag{48}$$

Finally, it can easily be shown that $\left(Y_m^*\right)^T \left(W_m^*\right)^T = KM\ R_m^*$, so that (48) becomes

$$A_m = KM\ R_m^* R_m \tag{49}$$

which is Parseval's theorem in view of (29) and because the elements of A_m are

$$a_{j,k} = M \sum_{s=0}^{K-1} P_{j,|m|}(\cos\theta_s)\ P_{k,|m|}(\cos\theta_s) \tag{50}$$

One can start with (49), being Parseval's Theorem, and show, first, that since A_m is a full rank matrix, R_m is non-singular; and, second, that solutions (31) and (44) are equivalent. Thus, there is a perfect duality between these solutions: (44) operates directly on the signal, g; (31) operates on its Fourier transform $\mathring{G}$; both yield identical results.

Harmonic Analysis of Averages

In order to make use of all gravity data around the world and to keep the computation of spherical harmonic models tractable, usual practice is to transform the data into mean quantities, i.e., averages over certain defined blocks or areas on the Earth's surface, often delineated by constant lines of latitude and longitude, e.g., 1°x1° or 30'x30' blocks. This yields a new function, $\bar{g}$, in principle, defined *for every point* on the sphere, but sampled, like g, on a regular grid. $\bar{g}$ has its own unique Legendre spectrum, $\{\bar{\gamma}_{n,m}\}$, and all the previous results hold with respect to this case. The question, however, is how to relate $\bar{\gamma}_{n,m}$ and $\gamma_{n,m}$, the latter being ultimately sought.

This relationship depends on the spherical area over which the averaging is performed. It is derived for the uniformly weighted angular block averages (e.g., 1°x1° mean anomalies) by Gaposchkin (1980). The relationship for the generally weighted angular block average is deduced as follows. Let the average be defined in terms of a convolution with the general weighting function, b:

$$\overline{g}(\theta,\lambda) = \frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi} b(\theta,\lambda,\theta',\lambda')\, g(\theta',\lambda')\, \sin\theta'\, d\theta'\, d\lambda' \tag{51}$$

And, suppose $B_{k,m}$ is the frequency response of the weighting function (the filter) *in terms of the Cartesian variables* θ,λ. Then by the convolution theorem for periodic functions on the plane

$$\overline{G}_{k,m} = B_{k,m}\, G_{k,m} \tag{52}$$

Using (19) for both $\overline{G}_{k,m}$ and $G_{k,m}$ yields

$$\sum_{n=|m|}^{\infty} \overline{\gamma}_{n,m}\, \rho_k^{n,|m|} = B_{k,m} \sum_{n=|m|}^{\infty} \gamma_{n,m}\, \rho_k^{n,|m|} \tag{53}$$

Now multiply both sides by $\sigma_{\cdot k}^{j,m}$, sum over k, and use the orthogonality (15) to get

$$\overline{\gamma}_{j,m} = \frac{\pi}{2\varepsilon_m} \sum_{n=|m|}^{\infty} \left(\sum_{k=-\infty}^{\infty} B_{k,m}\, \rho_k^{n,|m|}\, \sigma_{\cdot k}^{j,|m|} \right) \gamma_{n,m} \tag{54}$$

For purposes of discussion, suppose the filter, b, removes all power beyond the Nyquist frequencies on the plane, so that $B_{k,m} = 0$ for $|k| \geq K/2$ or $|m| \geq M/2$. (The response for the uniformly weighted block average is rather inefficient in this respect.) Then, clearly, $\overline{\gamma}_{j,m} = 0$ for $|m| \geq M/2$; but $\overline{\gamma}_{j,m} \neq 0$ for *any* degree, j, when $|m| < M/2$. In terms of the previous notation (see Figures 1 and 2), $\Delta\overline{\gamma}_m = 0$, but $\delta\overline{\gamma}_m \neq 0$. This means that there remains considerable aliasing of the spectrum estimated by any method of harmonic analysis if the averaging is performed over constant angular blocks. In fact, with a perfect angular block average (in addition to being zero at frequencies higher than the Nyquist, $B_{k,m} = 1$ for $|k| < K/2$ and $|m| < M/2$), equation (54) with (15) becomes $\overline{\gamma}_{j,m} = \gamma_{j,m}$ for any j if $|m| < M/2$. That is, the spherical spectrum denoted by the region $\delta\tilde{\gamma}$ in Figure 2 is passed fully by this filter and therefore represents an aliasing error (see equation 42).

A better filter is the so-called spherical cap average (Jekeli 1981). In this case (for constant size caps, only!), the frequency response of the filter is best formulated in terms of its Legendre spectrum, β_n; and

$$\overline{\gamma}_{n,m} = \beta_n\, \gamma_{n,m} \tag{55}$$

For the uniformly weighted average over a spherical cap of radius, ψ_c,

$$b(\psi) = \begin{cases} \dfrac{2}{1-\cos\psi_c}; & \psi \leq \psi_c \\ 0; & \text{otherwise} \end{cases} \tag{56}$$

where $\cos\psi = \cos\theta\cos\theta' + \sin\theta\sin\theta'\cos(\lambda - \lambda')$. The filter response, β_n, is the familiar Pellinen smoothing factor (Sjöberg 1980). Figure 3 shows β_n for $\psi_c = 0.5°$.

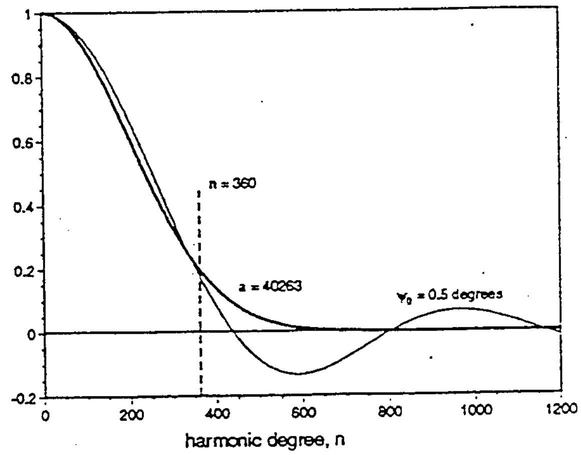


Fig.3: Pellinen smoothing factor (thin line) versus Gaussian smoothing factor (thick line).

If the filter were designed such that $\beta_n = 0$ for $n \geq K$ then from (55), clearly $\overline{\gamma}_{n,m} = 0$ for $n \geq K$ (i.e., both $\Delta\overline{\gamma}_m = 0$ *and* $\delta\overline{\gamma}_m = 0$); and the Legendre spectrum of the averaged function estimated by (31), (41), or (44) is not aliased. Such a filter is approximately given by the Gaussian weighting function (Jekeli 1981):

$$b(\psi) = e^{-a(1-\cos\psi)}; \quad a > 0;\ 0 \leq \psi \leq \pi \tag{57}$$

The corresponding response is also shown in Figure 3 for a = 40263. Both responses were designed to attenuate all harmonics with n > 360 to less than 20%. (The filter (51) with (57), being a *global* average, is approximately a cap average since $b(\psi) \approx 0$, $\psi \geq \psi_c(a)$.)

It is seen from (54) and (55) that the frequency responses of the constant angular block average, $B_{k,m}$, and of the cap average, β_n, are completely different, the first capable of removing $\Delta\gamma$, only, and the second able to remove both $\Delta\gamma$ and $\delta\gamma_m$. Therefore, also the practice of approximating the response of the constant angular block average by β_n is quite incorrect.

Data Errors and Least-Squares Collocation

So far, the data measurement noise and an a priori spectrum with weights have not been considered. These are important aspects of modern harmonic analyses. However, the computational tractability of the analysis for high-degree models (such as N = 360) restricts the allowable variances and covariances among the errors of the given data, as well as the weights of the a priori harmonic coefficients (Colombo 1981). In essence, the orthogonality of the matrices $\tilde{Y}_m$ (see 41) must be availed

to reduce the dimensionality of the normal matrices to be inverted.

If (38) is written as

$$g = \widetilde{Y}\,\widetilde{\gamma} \tag{58}$$

where $\widetilde{Y} = \left[\widetilde{Y}_{-N+1}, \ldots, \widetilde{Y}_{N-1}\right]$ and $\widetilde{\gamma} = \left[\widetilde{\gamma}^{T}_{-N+1}, \ldots, \widetilde{\gamma}^{T}_{N-1}\right]^{T}$, then the least-squares solution with weights based on the noise variance/covariance matrix, P^{-1}, is given by

$$\hat{\widetilde{\gamma}} = \left(\widetilde{Y}^{T} P\, \widetilde{Y}\right)^{-1} \widetilde{Y}^{T} P\, g \tag{59}$$

In order to take advantage of the orthogonality (40), it is required that

$$\left(\widetilde{Y}^{T}_{m} P\, \widetilde{Y}_{l}\right) = 0\,; \quad \text{if } m \neq l \tag{60}$$

This happens only if the errors in the data $\{g_{s,t}\}$ are uncorrelated in longitude, and if the correlation in latitude is the same for all longitudes.

The absence of redundancy of data in the new approach obviates weighting the data, and no restrictions need be imposed on the correlations in the data noise for either formulation, (44) or (31). The variance/covariance matrix of the estimated coefficients is obtained by simple error propagation.

Colombo's (1981) least-squares *collocation* (l.s.c) method of harmonic analysis, in theory, is supposed to yield the best estimator of the Legendre spectrum for a given set of data. In case the data (again, globally and regularly distributed on a grid) are perfect measurements (no measurement noise), l.s.c. minimizes the aliasing error. The l.s.c. estimator is also of the quadratures type, as defined by Colombo (1981, p.43, eq.2.60):

$$\hat{\gamma}_{n,m} = \sum_{s=0}^{K-1} \sum_{t=0}^{M-1} \eta^{s}_{n,m}\, g_{s,t}\, e^{-i2\pi tm/M} \tag{61}$$

where $\eta^{s}_{n,m}$ are "weights" to be determined by inverting a KxK matrix of discrete spectral components of the Toeplitz/circulant autocovariance matrix of the data. Although originally formulated to estimate only coefficients up to degree and order N, it is easily extended to estimate any number of coefficients, such as those in (28). Again, the formulation (61) is preserved in the presence of data noise only if the noise covariance matrix has a structure identical to the physical autocovariance matrix of the data.

Considering (37) and (45), it is readily seen that all the other methods discussed here, namely (22), (31), (41), and (44) are of the quadratures type, but with different weights; and those in (61) are the best in terms of minimizing the aliasing error *for a particular set of estimated coefficients*. But the optimality of (61) is based on the assumed isotropy and homogeneity of the (physical) covariance function of the data. These assumptions are quite strong and do not hold uniformly for all frequencies. Furthermore, the covariance function usually is based on a model with questionable fidelity at the high frequencies. Therefore, in practical applications l.s.c. (for example, Colombo 1981; Gleason 1987) may not be significantly superior to the estimation based on the harmonic analysis (31) or (44) with appropriate filtering, e.g., with (57), because aliasing thus can be well controlled and/or eliminated.

Summary

The methods of harmonic analysis on the sphere were investigated from the point of view of aliasing. The aliasing error was formulated in terms of spherical harmonic coefficients in the case when the function is sampled on a regular grid of latitudes and longitudes. This yielded several results, some of which are known, but perhaps more analytically illustrated, here: 1) The simple quadratures method and related methods are biased even with band-limited functions. 2) A new method that eliminates this bias is also superior to Colombo's method of least squares in terms of reducing aliasing, because it estimates more coefficients at little added computational cost. 3) But, a simple modification of the model makes the least squares method identical to the new method as one is the dual of the other. The new method solves for as many coefficients as data and, in the presence of data noise, it need not use fabricated noise covariances to make the solution numerically tractable. 4) The essential elimination of aliasing can only be effected with spherical cap averages, not with the often used constant angular block averages.

Appendix - The Computation of $\mathring{\rho}^{n,m}_{k}$

Analogous to (2b)

$$\mathring{\rho}^{n,m}_{k} = \frac{1}{K} \sum_{s=0}^{K-1} P_{n,m}(\cos\theta_s)\, e^{-i2\pi ks/K}\,, \quad k = 0, \ldots, K-1 \tag{A.1}$$

Also define the Discrete Fourier Transforms of cosθ and sinθ on the interval (0,π):

$$\mathring{\xi}_{k} = \frac{1}{K} \sum_{s=0}^{K-1} \cos\theta_s\, e^{-i2\pi ks/K}\,, \quad \mathring{\xi}_{k+K} = \mathring{\xi}_{k} \tag{A.2a}$$

$$\mathring{\zeta}_{k} = \frac{1}{K} \sum_{s=0}^{K-1} \sin\theta_s\, e^{-i2\pi ks/K}\,, \quad \mathring{\zeta}_{k+K} = \mathring{\zeta}_{k} \tag{A.2b}$$

The recursion formulas for the normalized associated

Legendre functions are

$$P_{n,m}(y) = c_{n-1,m}\, y\, P_{n-1,m}(y) - d_{n-2,m}\, P_{n-2,m}(y)\,, \quad 0 \le m \le n-2 \qquad \text{(A.3a)}$$

$$P_{n,n-1}(y) = \sqrt{2n+1}\; y\, P_{n-1,n-1}(y)\,, \quad n \ge 1 \qquad \text{(A.3b)}$$

$$P_{n,n}(y) = \sqrt{\frac{2n+1}{2n}}\, \sqrt{1-y^2}\; P_{n-1,n-1}(y)\,, \quad n \ge 2 \qquad \text{(A.3c)}$$

$$P_{0,0}(y) = 1\,, \quad P_{1,1}(y) = \sqrt{3}\, \sqrt{1-y^2} \qquad \text{(A.3d)}$$

where

$$c_{n-1,m} = \sqrt{\frac{(2n-1)(2n+1)}{(n-m)(n+m)}}\,,$$

$$d_{n-2,m} = \sqrt{\frac{(2n+1)(n+m-1)(n-m-1)}{(2n-3)(n-m)(n+m)}} \qquad \text{(A.3e)}$$

Substituting (A.3a-d) into (A.1) and using the convolution theorem for the product $\cos\theta_s\, P_{n,m}(\cos\theta_s)$ then yields recursion formulas for the $\mathring{\rho}_k^{n,m}$, $k = 0, \ldots, K-1$:

$$\mathring{\rho}_k^{n,m} = c_{n-1,m} \sum_{j=0}^{K-1} \mathring{\rho}_j^{n-1,m}\, \mathring{\zeta}_{j-k} - d_{n-2,m}\, \mathring{\rho}_k^{n-2,m}\,, \quad 0 \le m \le n-2 \qquad \text{(A.4a)}$$

$$\mathring{\rho}_k^{n,n-1} = \sqrt{2n+1} \sum_{j=0}^{K-1} \mathring{\rho}_j^{n-1,n-1}\, \mathring{\zeta}_{j-k}\,, \quad n \ge 1 \qquad \text{(A.4b)}$$

$$\mathring{\rho}_k^{n,n} = \sqrt{\frac{2n+1}{2n}} \sum_{j=0}^{K-1} \mathring{\rho}_j^{n-1,n-1}\, \mathring{\zeta}_{j-k}\,, \quad n \ge 2 \qquad \text{(A.4c)}$$

$$\mathring{\rho}_k^{0,0} = \delta(k,0)\,, \quad \mathring{\rho}_k^{1,1} = \sqrt{3}\, \mathring{\zeta}_k \qquad \text{(A.4d)}$$

Acknowledgments. This work was supported by contract F19628-93-K-0033 with the Air Force Phillips Laboratory, Hanscom AFB, MA. The author is grateful to the reviewers whose valuable comments led to many improvements.

References

Abramowitz M, Stegun IA (eds.) (1972) Handbook of Mathematical Functions. Dover Publications.

Albertella A, Sacerdote F (1995) Spectral analysis of block averaged data in geopotential global model determination. To appear in Manuscripta Geodaetica.

Albertella A, Sacerdote F, Sansò F (1993) Geodetic calculus with block-averages observations on the sphere. Surveys in Geophysics 14: 395–402.

Colombo OL (1981) Numerical methods for harmonic analysis on the sphere. Report no.310, Department of Geodetic Science, The Ohio State University.

Gaposchkin EM (1980) Averaging on the surface of a sphere. J Geophys Res 85(B6):3187-3193.

Gleason DM (1987) Developing an optimally estimated earth gravity model to degree and order 360 from a global set of 30' by 30' mean surface gravity anomalies. Manuscripta Geodaetica 12:253-267.

Goldstein JD (1978) Application of Fourier techniques to the computation of spherical harmonic coefficients. The Analytic Sciences Corp., Tech. Info. Mem. TIM-868-3.

Heiskanen WA, Moritz H (1967) Physical Geodesy. W.H. Freeman, San Francisco.

Hobson EW (1965) The Theory of Spherical and Ellipsoidal Harmonics. Chelsea Publishing.

Hwang C (1993) Fast algorithm for the formation of normal equations in a least-squares spherical harmonic analysis by FFT. Manuscripta Geodaetica 18(1):46-52.

Jekeli C (1981) Alternative methods to smooth the earth's gravity field. Report No.327, Department of Geodetic Science and Surveying, Ohio State University.

Kaula WM (1993) Higher Order Statistics of Planetary Gravities and Topographies. Geophys Res Lett 20.

Pavlis NK (1988) Modeling and estimation of a low degree geopotential model from terrestrial gravity data. Report No.386, Department of Geodetic Science and Surveying, The Ohio State University.

Priestley MB (1981) Spectral Analysis and Time Series. Academic Press.

Rapp RH (1977) Determination of Potential Coefficients to Degree 52 from 5° mean gravity anomalies. Bulletin Géodésique 51(4):301-323.

Rapp RH, Pavlis NK (1990) The development and analysis of geopotential coefficient models to spherical harmonic degree 360. J Geophys Res 95(B13):21885-21911.

Ricardi LJ, Burrows ML (1972) A recurrence technique for expanding a function in spherical harmonics. IEEE Trans Comp June:583-585.

Schmitz DR, Cain JC (1983) Geomagnetic Spherical Harmonic Analysis, 1. Techniques. J Geophys Res 88(B2):1222-1228.

Sansò F (1990) On the aliasing problem in the spherical harmonic analysis. Bulletin Géodésique 64:313-330.

Sjöberg L (1980) A recurrence relation for the β_n-function. Bulletin Géodésique 54(1):69-72.